

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 25, 2005, 15:31:28 ; Search time 2861 Seconds
(without alignments)
762.141 Million cell updates/sec

Title: US-10-089-278-6_COPY_1_45

Perfect score: 229
Sequence: 1 MCLTLNLSAEGKVDQSKI.....AMWGFOTTAETVSTARAQPA 45

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh
-Q=/cpn2_1/USPTO.spool_h/US10089278/runat_25082005_131342_29645/app_query.fasta_1.199
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10089278 @CGN 1.1 3731 @runat_25082005_131342_29645 -NCFU=6 -ICPU=3
-NO MWAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	229	100.0	906	6	AX111698 Sequence
2	229	100.0	927	6	AX111695 Sequence
3	229	100.0	946	6	AX111697 Sequence
4	229	100.0	990	6	AX111696 Sequence

5	229	100.0	1030	6	AX111694
6	200	87.3	4776	6	A95117
7	200	87.3	4776	6	A95152
8	200	87.3	4776	6	AR304364
9	191	83.4	220	6	AR034255
10	191	83.4	220	6	BD235640
11	191	83.4	220	6	AR287725
12	183	79.9	1704	14	SNVENVORF
13	170	74.2	8221	14	AF246698
14	161	70.3	3149	14	REVARELE
15	108.5	47.4	9602	6	AX573109
16	108.5	47.4	9602	6	AX573110
17	104	45.4	5082	6	AX573112
18	67	29.3	597	10	MUSIGHAPB
19	66.5	29.0	1515	8	AK111428
20	66.5	29.0	132703	8	CNS08CBT
21	66.5	29.0	140709	9	HS103M22
22	66	28.8	88203	5	AC097628
23	66	28.8	188104	2	AC094551
24	65.5	28.6	620	5	CPU50606
25	65.5	28.6	7677	8	AF178967
26	65.5	28.6	24479	3	AC114265
27	65	28.4	110000	2	AC096315_6
28	65	28.4	110000	2	AC096315_7
29	65	28.4	165990	2	AC142089
30	65	28.4	225103	2	AC137199
31	65	28.4	233024	2	AC103343
32	64	27.9	170943	10	AL596095
33	64	27.9	198614	2	AC073293
34	64	27.9	213606	2	AC023811
35	64	27.9	229312	10	AC134603
36	63.5	27.7	14234	9	AF133901
37	63	27.5	978	6	AX434448
38	63	27.5	110000	1	AE017333_30
39	63	27.5	110000	1	CP000002_30
40	63	27.5	180852	2	CR762436
41	63	27.5	244856	2	AC094026
42	62.5	27.3	3067	10	BC024874
43	62.5	27.3	37061	9	AP001064
44	62.5	27.3	110000	2	AC106166_2
45	62.5	27.3	115798	9	AP001065

ALIGNMENTS

RESULT 1	AX111698	AX111698	906 bp	DNA	linear	PAT 30-APR-2001
LOCUS	Sequence 5 from Patent WO0125415.					
DEFINITION	AX111698					
ACCESSION	AX111698.1	GI:13927963				
VERSION						
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	other sequences; artificial sequences.					
REFERENCE	1					
AUTHORS	Cichutek, K. and Engelstaedter, M.					
TITLE	Gene transfer in human lymphocytes using retroviral scfv cell targeting					
JOURNAL	Patent: WO 0125415-A 5 12-APR-2001;					
FEATURES	Bundesrepublik Deutschland LER (DE)					
source	Location/Qualifiers					
	1..906					
	/organism="synthetic construct"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:32630"					
	/note="scFv kodierende Sequenz"					

ORIGIN	Alignment Scores:	6,278-22	Length:	906
	Pred. No.:	229.00	Matches:	45
	Score:	100.00%	Conservative:	0

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x AX111698 (1-906)

QY 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyValAspGlnAlaSerLysIle 20
Db 1 ATGGACTGTCTCACCACCTCCGATCCGTGAGGGTAAAGTTGACCGAGCGAGCAAAATC 60
QY 21 LeuIleLeuValAlaTrpTrpGlyPheGlyThrAlaGluValSerThrAlaArg 40
Db 61 CTAATTCTCCTTGTGGCTTGGTGGGGTTTGGGACCACCTGCCGAAGTTTCGACTGCCCGA 120
QY 41 AlaAlaGlnProAla 45
Db 121 GCGGCCAGCGCGCC 135

RESULT 2
LOCUS AX111695 927 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 2 from Patent WO0125415.
ACCESSION AX111695
VERSION AX111695.1 GI:13927960
KEYWORDS
SOURCE
ORGANISM
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Cichutek,K. and Engelstaedter,M.
TITLE Gene transfer in human lymphocytes using retroviral scfv cell
targeting
JOURNAL
Patent: WO 0125415-A 2 12-APR-2001;
Bundesrepublik Deutschland LET (DE)
FEATURES
Location/Qualifiers
1..946
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="scFv kodierende Sequenz"

ORIGIN
Alignment Scores:
Pred. No.: 6.56e-22 Length: 946
Score: 229.00 Matches: 45
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x AX111697 (1-946)

QY 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyValAspGlnAlaSerLysIle 20
Db 1 ATGGACTGTCTCACCACCTCCGATCCGTGAGGGTAAAGTTGACCGAGCGAGCAAAATC 60
QY 21 LeuIleLeuValAlaTrpTrpGlyPheGlyThrAlaGluValSerThrAlaArg 40
Db 61 CTAATTCTCCTTGTGGCTTGGTGGGGTTTGGGACCACCTGCCGAAGTTTCGACTGCCCGA 120
QY 41 AlaAlaGlnProAla 45
Db 121 GCGGCCAGCGCGCC 135

RESULT 4
LOCUS AX111696 990 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 3 from Patent WO0125415.
ACCESSION AX111696
VERSION AX111696.1 GI:13927961
KEYWORDS
SOURCE
ORGANISM
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Cichutek,K. and Engelstaedter,M.
TITLE Gene transfer in human lymphocytes using retroviral scfv cell
targeting
JOURNAL
Patent: WO 0125415-A 3 12-APR-2001;
Bundesrepublik Deutschland LET (DE)
FEATURES
Location/Qualifiers
1..990
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="scFv kodierende Sequenz"

ORIGIN
Alignment Scores:
Pred. No.: 6.87e-22 Length: 990
Score: 229.00 Matches: 45
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x AX111696 (1-990)

QY 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyValAspGlnAlaSerLysIle 20

```

Db 1 ATGGACTGTCTCACCACCTCCGATCCGCTGAGGGTAAAGTTGACACGGCGCAAAATC 60
Qy 21 LeuileLeuValAlaTrrpGlyPheGlyThrAlaGluValSerThrAlaArg 40
Db 61 CTAATTCTCTTGTGGCTTGGTGGGGTTGGGACCACTGCCGAAGTTTCGACTGCCCGA 120
Qy 41 AlaAlaGlnProIla 45
Db 121 GCGGCCAGCCGCC 135
RESULT 5
AX111694
LOCUS AX111694 1030 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1 from Patent WO0125415.
ACCESSION AX111694
VERSION AX111694.1 GI:13927959
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Cichutek,K. and Engelstaedter,M.
TITLE Gene transfer in human lymphocytes using retroviral scfv cell targeting
JOURNAL Patent: WO 0125415-A 1 12-APR-2001;
Bundesrepublik Deutschland LET (DE)
FEATURES
source 1. .1030
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="scfv kodierende Sequenz"
ORIGIN
Alignment Scores:
Pred. No.: 7.15e-22 Length: 1030
Score: 229.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-089-278-6_COPY_1_45 (1-45) x AX111694 (1-1030)
Qy 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyValAspGlnAlaSerLysIle 20
Db 44 ATGGACTGTCTCACCACCTCCGATCCGCTGAGGGTAAAGTTGACACGGCGCAAAATC 103
Qy 21 LeuileLeuValAlaTrrpGlyPheGlyThrAlaGluValSerThrAlaArg 40
Db 104 CTAATTCTCTTGTGGCTTGGTGGGGTTGGGACCACTGCCGAAGTTTCGACTGCCCGA 163
Qy 41 AlaAlaGlnProIla 45
Db 164 GCGGCCAGCCGCC 178
RESULT 6
A95117
LOCUS A95117 4776 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from Patent WO9928489.
ACCESSION A95117
VERSION A95117.1 GI:6779272
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1
AUTHORS Cichutek,K. and Engelstaedter,M.
TITLE Cell-specific retroviral vectors with antibody domains and method for the production thereof for selective gene transfer
JOURNAL Patent: WO 9928489-A 1 10-JUN-1999;
CICHUTEK KLAUS (DE); BUNDESREPUBLIK DEUTSCHLAND LET (DE); ENGELSTAEDTER MARTIN (DE)

FEATURES
source 1. .4776
Location/Qualifiers
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Alignment Scores:
Pred. No.: 3.36e-17 Length: 4776
Score: 200.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 87.34% Indels: 0
DB: 6 Gaps: 0
US-10-089-278-6_COPY_1_45 (1-45) x A95117 (1-4776)
Qy 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyValAspGlnAlaSerLysIle 20
Db 1025 ATGGACTGTCTCACCACCTCCGATCCGCTGAGGGTAAAGTTGACACGGCGCAAAATC 1084
Qy 21 LeuileLeuValAlaTrrpGlyPheGlyThrAlaGluValSerThrAla 39
Db 1085 CTAATTCTCTTGTGGCTTGGTGGGGTTGGGACCACTGCCGAAGTTTCGACTGCC 1141
RESULT 7
A95152
LOCUS A95152 4776 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from Patent WO9928488.
ACCESSION A95152
VERSION A95152.1 GI:6779277
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1
AUTHORS Cichutek,K. and Merget-Millitzer,H.
TITLE Pseudo-type retroviral vectors with modifiable surface capsid proteins
JOURNAL Patent: WO 9928488-A 1 10-JUN-1999;
CICHUTEK KLAUS (DE); BUNDESREPUBLIK DEUTSCHLAND LET (DE); MERGET MILLITZER HEIKE (DE)
FEATURES
source 1. .4776
Location/Qualifiers
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Alignment Scores:
Pred. No.: 3.36e-17 Length: 4776
Score: 200.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 87.34% Indels: 0
DB: 6 Gaps: 0
US-10-089-278-6_COPY_1_45 (1-45) x A95152 (1-4776)
Qy 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyValAspGlnAlaSerLysIle 20
Db 1025 ATGGACTGTCTCACCACCTCCGATCCGCTGAGGGTAAAGTTGACACGGCGCAAAATC 1084
Qy 21 LeuileLeuValAlaTrrpGlyPheGlyThrAlaGluValSerThrAla 39
Db 1085 CTAATTCTCTTGTGGCTTGGTGGGGTTGGGACCACTGCCGAAGTTTCGACTGCC 1141
RESULT 8
AR304364
LOCUS AR304364 4776 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6544779.
ACCESSION AR304364
VERSION AR304364.1 GI:31693481

```

KEYWORDS      .
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE      1 (bases 1 to 4776)
AUTHORS       Cichutek, K. and Merget-Millitzer, H.
TITLE         Pseudo-type retroviral vectors with modifiable surface capsid
JOURNAL       Patent: US 6544779-A 1 08-APR-2003;
FEATURES      Location/Qualifiers
               source
               1..4776
               /organism="unknown"
               /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      3..36e-17      Length:      4776
Score:          200.00         Matches:    39
Percent Similarity: 100.00%     Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:    87.34%         Indels:     0
DB:             6              Gaps:       0

US-10-089-278-6_COPY_1_45 (1-45) x AR304364 (1-4776)
Qy      1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyValAspGlnAlaSerLysIle 20
Db      1025 ATGGACTGTCTCACCACCTCCGATCCGCTGAGGGTAAAGTTGACCAGCGGAGCAAAATC 1084
Qy      21 LeuIleLeuValAlaIleTrpGlyPheGlyThrAlaGluValSerThrAla 39
Db      1085 CTAATTCTCTTGTGGCTTGGTGGGGTTTGGACCACTGCCGAGCTTCGACTGCC 1141

RESULT 9
AR034255
LOCUS      AR034255
DEFINITION Sequence 2 from patent US 5869331.
ACCESSION AR034255
VERSION    AR034255.1 GI:5949860
KEYWORDS  .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 220)
AUTHORS    Dornburg, R.C.
TITLE      Cell type specific gene transfer using retroviral vectors
           containing antibody-envelope fusion proteins and wild-type envelope
JOURNAL    Patent: US 5869331-A 2 09-FEB-1999;
FEATURES   Location/Qualifiers
           source
           1..220
           /organism="unknown"
           /mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.:      2.5e-17      Length:      220
Score:          191.00         Matches:    37
Percent Similarity: 100.00%     Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:    83.41%         Indels:     0
DB:             6              Gaps:       0

US-10-089-278-6_COPY_1_45 (1-45) x AR034255 (1-220)
Qy      1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIle 20
Db      85  ATGGACTGTCTCACCACCTCCGATCCGCTGAGGGTAAAGTTGACCAGCGGAGCAAAATC 144
Qy      21 LeuIleLeuValAlaIleTrpGlyPheGlyThrAlaGluValSer 37
Db      145 CTAATTCTCTTGTGGCTTGGTGGGGTTTGGACCACTGCCGAGCTTCG 195

KEYWORDS      .
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE      1 (bases 1 to 4776)
AUTHORS       Cichutek, K. and Merget-Millitzer, H.
TITLE         Pseudo-type retroviral vectors with modifiable surface capsid
JOURNAL       Patent: US 6544779-A 1 08-APR-2003;
FEATURES      Location/Qualifiers
               source
               1..4776
               /organism="unknown"
               /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      3..36e-17      Length:      4776
Score:          200.00         Matches:    39
Percent Similarity: 100.00%     Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:    87.34%         Indels:     0
DB:             6              Gaps:       0

US-10-089-278-6_COPY_1_45 (1-45) x AR304364 (1-4776)
Qy      1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyValAspGlnAlaSerLysIle 20
Db      1025 ATGGACTGTCTCACCACCTCCGATCCGCTGAGGGTAAAGTTGACCAGCGGAGCAAAATC 1084
Qy      21 LeuIleLeuValAlaIleTrpGlyPheGlyThrAlaGluValSerThrAla 39
Db      1085 CTAATTCTCTTGTGGCTTGGTGGGGTTTGGACCACTGCCGAGCTTCGACTGCC 1141

```

```

RESULT 10
BD235640
LOCUS      BD235640
DEFINITION Cell type-specific gene transfer with the use of retrovirus vector
           having antibody envelope-fused protein and wild type envelope
           protein.
ACCESSION BD235640
VERSION    BD235640.1 GI:33045410
KEYWORDS   JP 2002522090-A/4.
SOURCE     synthetic construct
           ORGANISM
           1 (bases 1 to 220)
           other sequences; artificial sequences.
REFERENCE  1 (bases 1 to 220)
AUTHORS    Dornburg, R.C.
TITLE      Cell type-specific gene transfer with the use of retrovirus vector
           having antibody envelope-fused protein and wild type envelope
           Patent: JP 2002522090-A 4 23-JUL-2002;
           THOMAS JEFFERSON UNIVERSITY
           OS Artificial Sequence
           PN JP 2002522090-A/4
           PD 23-JUL-2002
           PF 10-AUG-1999 JP 2000565164
           PR 17-AUG-1998 US 09/135121
           PI RALPH C DORNBURG
           PC C12N15/09,A61K48/00,C12N7/00//A61K35/76,C07K14/16,C07K16/28,
           C07K19/00,
           PC (C12N7/00,C12R1:92),C12N15/00
           PC (C12N7/00,C12R1:92),C12N15/00
           CC pTC13 eucaryotic gene expression vector
           FH Key Location/Qualifiers
           FT source 1..220
           FT Location/Qualifiers
           /organism='Artificial Sequence'.
           /mol_type='synthetic construct'
           /db_xref='taxon:32630'

ORIGIN
Alignment Scores:
Pred. No.:      2.5e-17      Length:      220
Score:          191.00         Matches:    37
Percent Similarity: 100.00%     Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:    83.41%         Indels:     0
DB:             6              Gaps:       0

US-10-089-278-6_COPY_1_45 (1-45) x BD235640 (1-220)
Qy      1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIle 20
Db      85  ATGGACTGTCTCACCACCTCCGATCCGCTGAGGGTAAAGTTGACCAGCGGAGCAAAATC 144
Qy      21 LeuIleLeuValAlaIleTrpGlyPheGlyThrAlaGluValSer 37
Db      145 CTAATTCTCTTGTGGCTTGGTGGGGTTTGGACCACTGCCGAGCTTCG 195

RESULT 11
AR287725
LOCUS      AR287725
DEFINITION Sequence 4 from patent US 6534051.
ACCESSION AR287725
VERSION    AR287725.1 GI:31674729
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 220)
AUTHORS    Dornburg, R.
TITLE      Cell type specific gene transfer using retroviral vectors
           containing antibody-envelope fusion proteins
           Patent: US 6534051-A 4 18-MAR-2003;
           Location/Qualifiers
           JOURNAL
           FEATURES

```



```

source 1. .220
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 2,5e-17 Length: 220
Score: 191.00 Matches: 37
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 83.41% Indels: 0
DB: Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x AR287725 (1-220)
QY 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysile 20
|||||
DB 85 ATGGACTGTCTACCAACCTCCGATCCGCTGAGGTAAAGTTGACCGAGCGCAAAATC 144
|||||
QY 21 LeuileuLeuValAlaTrpTrpGlyPheGlyThrAlaGluValSer 37
|||||
DB 145 CTAATTCTCCTGTGGCTTGGTGGGGTTTGGGACCACTGCCGAGTTTCG 195
|||||

RESULT 12
SNVENVORF
LOCUS SNVENVORF 1704 bp DNA linear VRL 03-AUG-1993
DEFINITION Spleen necrosis virus (clone pPB101) envelope open reading frame
sequence.
ACCESSION M87666
VERSION M87666.1 GI:334986
KEYWORDS
SOURCE Spleen necrosis virus
ORGANISM Spleen necrosis virus
VIRUSES; Retroviral viruses; Retroviridae; Mammalian type C
retroviruses; 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 1704)
AUTHORS Kewalramani,V.N., Panganiban,A.T. and Emerman,M.
TITLE Spleen necrosis virus, an avian immunosuppressive retrovirus,
shares a receptor with the type D simian retroviruses
JOURNAL J. Virol. 66 (5), 3026-3031 (1992)
MEDLINE 92219390
PUBMED 1313915
COMMENT Original source text: Spleen necrosis virus DNA.
FEATURES
source
1. .1704
Location/Qualifiers
/organism="Spleen necrosis virus"
/proviral
/mol_type="genomic DNA"
/db_xref="taxon:11836"

ORIGIN
Alignment Scores:
Pred. No.: 2,56e-15 Length: 1704
Score: 183.00 Matches: 35
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.91% Indels: 0
DB: Gaps: 14

US-10-089-278-6_COPY_1_45 (1-45) x SNVENVORF (1-1704)
QY 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysile 20
|||||
DB 1 ATGGACTGTCTACCAACCTCCGATCCGCTGAGGTAAAGTTGACCGAGCGCAAAATC 60
|||||
QY 21 LeuileuLeuValAlaTrpTrpGlyPheGlyThrAlaGlu 35
|||||
DB 61 CTAATTCTCCTGTGGCTTGGTGGGGTTTGGGACCACTGCCGAA 105
|||||

RESULT 13
AF246698
LOCUS AF246698 8221 bp DNA linear VRL 12-MAR-2003
DEFINITION Fowlpox virus hypothetical protein gene, partial cds; integrated

```

```

reticuloendotheliosis provirus, complete sequence; and protein
kinase-like protein gene, partial cds.
AF246698
AF246698.2 GI:28927668

Fowlpox virus
Fowlpox virus
VIRUSES; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
Avipoxvirus.
REFERENCE 1 (bases 1 to 8221)
AUTHORS Singh,P., Kim,T.-J. and Tripathy,D.N.
TITLE Re-emerging fowlpox: evaluation of isolates from vaccinated flocks
JOURNAL Avian Pathol. 29, 449-455 (2000)
REFERENCE 2 (bases 1 to 8221)
AUTHORS Singh,P., Schitzlein,W.M. and Tripathy,D.N.
TITLE Reticuloendotheliosis Virus Sequences Within the Genomes of Field
Strains of Fowlpox Virus Display Variability
J. Virol. (2003) In press
REFERENCE 3 (bases 1 to 8221)
AUTHORS Schmitzlein,W.M., Srinivasan,V. and Tripathy,D.N.
TITLE Direct Submission
JOURNAL Submitted (16-MAR-2000) Veterinary Pathobiology, University of
Illinois, 2001 South Lincoln Avenue, Urbana, IL 61802, USA
REFERENCE 4 (bases 1 to 8221)
AUTHORS Schmitzlein,W.M., Singh,P., Srinivasan,V. and Tripathy,D.N.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2003) University of Illinois at Urbana-Champaign,
2001 South Lincoln Avenue, Urbana, IL 61802, USA
REMARK Sequence update by submitter
COMMENT On Mar 12, 2003 this sequence version replaced gi:8926141.
FEATURES
source
1. .8221
Location/Qualifiers
/organism="Fowlpox virus"
/virion
/mol_type="genomic DNA"
/db_xref="taxon:10261"
/ncbi="University of Illinois field strain"
65. .8053
/organism="Fowlpox virus"
/mol_type="genomic DNA"
/db_xref="taxon:10261"
/ncbi="endogenous virus: reticuloendotheliosis provirus;
University of Illinois field strain; integrated provirus"
<1. .60
/ncbi="FPV ORF 201"
/codon_start=1
/evidence=not experimental
/product="hypothetical protein"
/protein_id="AAO62317.1"
/db_xref="GI:28927671"
/translation="NNDTSLPVAVKVIYGTVTI"
65. .575
972. .2471
/codon_start=1
/product="gag protein"
/protein_id="AAO62318.1"
/db_xref="GI:28927672"
/translation="MGOAGSKLLTPLEILKNSDFKFKAGDYGVDSFTLRKLCE
LEWPTFGVGMKEGTLDFRVVAVERNVEFGNPGHPDOVIYITWIDITIERPKYLKDC
GCKPTGSKVLLASOKNPNRPVLPSPAPESPRIRRAQFLDERPLSPAPPPPYDEV
FAIAEEGEGQQLQASQVNSPPHRSGLFEGAGSGSMYPLKETGERDMGGRPMRTYV
PFTTSDLYNWKQNPSFSQAPDEVISLESVYFTTHPTWDDCQQLRTLFTTEERBRV
RTEGRVNDQGVQVTDIEIAQFPATRPDMDNTGRNDNLRYQLILGRRAA
ARKPNLSKITVEVQAGDESPATVRLRYQVATWSPIDPRAPENAAVIOFVSQA
PDIRKIOKIDGFGKSLSELVAIAQVDFDREDAKATHELTKWAKVLLAQESRAE
RGSKXTPGKGRPPPLGNKQACKEEGHKKNCPLVSGATPVLVEELQ"
972. .1310
/product="matrix protein"
/ncbi="pro/pol"
<2472. .6053
/ncbi="pro/pol"
/codon_start=1
/product="protease/polymerase"
/protein_id="AAF81697.2"

```

```

/db_xref="GI:28927669"
/translation="GROGSSALREPRLLKVKVGGQILDFLVDGTGATHSVVQKPVGPMK
ESVALIGTNRINPKSEGRVLDLGRVLVTHSFVIPCDDPLGRDLLQKLRATIS
FTGEGPPTERTKLLVTALEERYFLLEAPIONVTLLQWKREIPKVAEINPGL
ASTQAPVHQLSLTPVVRQYPTITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLL
PVKSGTSYRMVQDLREVNKVEITHLPVNPYTLTLLPDRTWYSVLDLKDAPFC
IPLAESQLIFAFENADAESEGGQLTWRLPQGFKNSTPLDEALNRDLOGFLRDP
FVSLQYVDDLLIADTQAACATRDLLMTLAELOGVIRVSGKKAQLCQREVLGFKI
HKSSTLSNRSTAOILQIPVKTKQREFLGTIGYCRUWIPGFAELAQPLVYATRG
NDPLVGEKEEAFQSKLALTQPPALPSLDKPFQLFVEETGGAAGVLTQALGPW
KRPVJYSKRLDPVAAQWPCRLATAAALTRFASKTFQDIIETISHNLSLRS
PDRMLTNAITQYOVLLDPVRVFKOTAAINPATLLPRTDDTLP IHHCLTDLSTL
STRPDLTDPLAQAEATLFTDSSGYIRDGKRYTGAANVTLDVSIWAELPIGTSQA
ELIALTKALEWSDKSNYITDSYAFATLHVGMTRYRGLTTAGGKAIKNAPILA
LITAVPLKRVAVHCKHQKDDAPTSTGNRRDAVEVAIRPLSTQATISDAPDM
DTETPOYSSEALHGRLRGTQDPAGMWHLPDGRLLPRAVGRKVLQETHRAHGES
KLTELVRKHYLCGYRAARDITTRCVAQACVPNGATPVEKGLNSRIRGAAGPHEV
DTEMTVAKGGYKYLVLVDTFSGWVEYAPARETSQVVIKHLIHDIIIPRFLGPVQIG
SNGPASAVALYQQLCEALNVSKLHCAYRPOSSQGVEMRNTKETAKLRIETGG
WVSLPQALRARCTPREGLSFFELLYGLKPPVVRVCGDKLASITNQLKLSQAL
QATRSLARALRDQLPQKEAQDRTPLFQGDVLFVKHDFQGLGRMDGPYTVVLS
PTAVKVAGKTPMIIHHSRLKAPDQEWTVSPTDPLRVKLTTRAKP"
5990. .7750
/notes="env"
/codon_start=1
/products="envelope glycoprotein"
/protein_id="AAF81698.2"
/db_xref="GI:28927670"
/translation="MDCITNLSRABGKVDQAGKTLILVVMWFGTTABGYPLQQLWG
LPCDCGQVVFSPITVYTNLSGGSSTAYLTVGSGTGMGCGFRQWECVFKPKII
PSVQGPQPCSECLTIATQMHSTCYEKAQECTLLGKTVFTAILQKTKLGSVEDGPNK
LLOASTGIRGPVCMVPAVYVSDGQPTDMIRSESVRELEIIRHSYPSVQVHP
LALPRGVLDLPQTDILEATHVNLATNPQALNCWMLCTGLTPIPAIPANGVY
LNGNCSLSPFRVQPTSDVNCYAGEADNRGTIGYVYFTNCTSIQEVSNESHIR
NUTRLCPGPHVFCVGNMAYTALPNKIGLCILASIVPDMISIIIGEEPIPLPSEY
AGRHKRAVOIIPLLVGLGTGATLAGGTGLSVYTHKLSNQLIEDVQALSGTNDL
OQDILSAEVLQNRGELDLTAEQGIICLAEQKCFVANKSGIVRDKIRKLQDIL
ARKRALYNPLNGLNGFLPYLLPLGLFLGLILFTLGLPCIKLTIRIHKIQAVK
ILALVPPQYKLPTEMDTLGQ"
7804. .8053
8124. .>8221
/notes="ORF 203"
/codon_start=1
/products="protein kinase-like protein"
/protein_id="AAO62319.1"
/db_xref="GI:28927673"
/translation="MDIITNTMTFIQNDIPNIPYVDIEKPLLVS"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 7.88e-13 Length: 8221
Score: 170.00 Matches: 32
Percent Similarity: 91.43% Conservative: 0
Best Local Similarity: 91.43% Mismatches: 3
Query Match: 74.24% Indels: 0
DB: 14 Gaps: 0

```

US-10-089-278-6_COPY_1_45 (1-45) x AF246698 (1-8221)

```

QY 1 MetAspCysLeuThrAenLeuAtrGserAlaGluGlyLysValAspGlnAlaSerLysIle 20
Db 5990 ATGACTGTCTACCAACCTCCGATCGCTGAGGTAAAGTTGACCGAGCGGCAAAACC 6049
QY 21 LeuIleLeuValAlaTrpTrpGlyPheGlyThrAlaGlu 35
Db 6050 CTAATCTCTTGTGGTTTGGTGGGGGTTTGGGACCACTGCCGAG 6094

```

RESULT 14

```

REVARELE
LOCUS REVARELE 3149 bp RNA linear VRL 10-JUL-1995
DEFINITION Reticuloendotheliosis virus strain A pol (3' end) and env genes.
ACCESSION X01455 K02537
VERSION X01455.1 GI:61784

```

```

KEYWORDS capsid protein; env gene; glycoprotein; long terminal repeat;
overlapping genes; pol gene; reverse transcriptase.
SOURCE Reticuloendotheliosis virus
ORGANISM Reticuloendotheliosis virus
REFERENCE 1. Wilhelmsen, K.C., Eggleston, K. and Temin, H.M.
AUTHORS Wilhelmsen, K.C., Eggleston, K. and Temin, H.M.
TITLE Nucleic acid sequences of the oncogene v-rel in
reticuloendotheliosis virus strain 1 and its cellular homolog, the
proto-oncogene c-rel
JOURNAL J. Virol. 52 (1), 172-182 (1984)
MEDLINE 85009850
PUBMED 6090694
COMMENT On Oct 22, 2003 this sequence version replaced gi:1311516.
*source strain=A;
Rev-A is the non-defective helper virus of Rev-T (see REVTVREL).
FEATURES
Location/Qualifiers
1..3149
/organism="Reticuloendotheliosis virus"
/mo_type="genomic RNA"
/db_xref="taxon:11636"
<1..1422
/codon_start=1
/product="C end of pol protein (reverse transcriptase)"
/protein_id="CAA25685.1"
/db_xref="GI:61785"
/db_xref="GOA:P03360"
/db_xref="UniProt/Swiss-Prot:P03360"
/translation="KGEITTAGKAIKNAPILALLITAVLWPKRVAVHCKHQKDDA
PTSGNRRADVAEVAIRPLSTQATISDAPMDPTETPOYSHVEALGHLRGTKDS
AGMHLDPGRLLPRAVGRKYLEOTHRATHLGESKLTETVRKHVPCIGIYRAARDITT
RCVAQYNPRAAPVEKGLNSRIRGAAGHEWVDFTMITAKGGYKYLVLVDITFSG
WEAYPAKRETSQVVIKHLIDIIIPRFLPQIGSDNGPFAVAKYQQLCEALNVSK
LHCAYRPOSSQGVEMRNTLKKATKEDRRGRGLPFPSPGAFPTVREGSSAPP
BILYGLKPPVVRVCGDKLASITNQLKLSQALQATRSLARAAARPTAPERSARP
PTVNLKPSFFVKHDFQGLGRMDGPYTVVLSPTAVKAAAGKTPWIHYSRLKKAAPD
QEWTVSPTDPLRVKLTTRAKP"
814. .822
/notes="pt. glycosylation site"
misc_feature 877. .885
misc_feature 1072. .1080
misc_feature 1359. .3107
CDS
/notes="(aa 1-582)"
/codon_start=1
/product="env-protein (capsid protein)"
/protein_id="CAA25686.1"
/db_xref="GI:61786"
/db_xref="GOA:P03399"
/db_xref="UniProt/Swiss-Prot:P03399"
/translation="MDCITDLSRTEGKVDQAGKTLILVVMWFGTTABGHPLOQLWE
LPCDCGQVVFSPDLPITTPSIASVAPSLDRLVLOGSMGCGFRQWECVFKPKII
PSVQGPQPCSECLTIATQMHSTCYEKAQECTLLGKTVFTAILQKTKLGSYEDGPNKL
QASCTGIWETSMDLGRPCVCLDGGGPTDRFGRI CAEGLEIIIRHSYPSVQVHPALAP
RPRGVLDLPQTDILEATHVNLATNPQALNCWMLCTGLTQSPQPSRRMAMLSNEI
AVLASGATHRVNRQCLLEADNRGTIGYVYVHTNCTSIQESIITRVIYEILRDY
VLHVRVLYCFOHATGATLPPNKGILCILASIVPDMISIIIGEEPIPLPSEYTAGRHR
AVQFTPLVGLGITGATLAGGTGLSVYTHKLSNQLIEDVQALSGTNDLQDQIDS
LAEVQLNRRGLDILLTAEQGIICLAEQKCFVANKSGIVRDKIRKLQEDLLARKAL
YDPLNGLNGFLPYLLPLGLFLGLILFTLGLPCIRKLTIRIHKIQGSKNPRISP
AVQATPNRDGVPKRVNV"
2259. .2267
/notes="pot. glycosylation site"
misc_feature 2298. .2306
misc_feature 2811. .2819
misc_feature
/notes="pot. glycosylation site"
misc_feature 3132. .3149
misc_feature
/notes="5' end of 3' long terminal repeat (LTR)"
misc_feature 3132. .3136
/notes="inverted repeat marking LTR"
ORIGIN

```

Alignment Scores:

Pred. No.:	5.11e-12	Length:	3149
Score:	161.00	Matches:	30
Percent Similarity:	88.57%	Conservative:	1
Best Local Similarity:	85.71%	Mismatches:	4
Query Match:	70.31%	Indels:	0
DB:	14	Gaps:	0

US-10-089-278-6_COPY_1_45 (1-45) x REVARELE (1-3149)

Qy	1	MetAspCysLeuThrAsnLeuArgSerAlaGluGlyValAspGlnAlaSerIysile	20
Db	1359	ATGGACTGTCACCCAGCTCCGATCCACTGAGGGTAAAGTTGACCAGGCGGCAAAACC	1418
Qy	21	LeuIleLeuLeuValAlaTrpTrpGlyPheGlyThrThrAlaGlu	35
Db	1419	CTAATTCCTCTCTGGTTGGTGGGGTTTGGACCACCTGCCGAG	1463

RESULT 15

AX573109	9602 bp	DNA	linear	PAT 29-NOV-2002
LOCUS				
DEFINITION	Sequence 4 from Patent WO02059338.			
ACCESSION	AX573109			
VERSION	AX573109.1	GI:26005048		

KEYWORDS

synthetic construct
synthetic construct
other sequences; artificial sequences.

ORGANISM

REFERENCE

1 Summerford, C., Gray, J.T., Lee, J.S. and Mulligan, R.C.
Retroviral vectors for transduction into quiescent cells and
packaging systems for them
Patent: WO 02059338-A 4 01-AUG-2002;
The Children's Medical Center Corporation (US) ; PRESIDENT AND
FELLOWS OF HARVARD COLLEGE (US)

FEATURES

source Location/Qualifiers

1..9602
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="retroviral vector"

ORIGIN

Alignment Scores:

Pred. No.:	0.000269	Length:	9602
Score:	108.50	Matches:	23
Percent Similarity:	67.57%	Conservative:	2
Best Local Similarity:	62.16%	Mismatches:	3
Query Match:	47.38%	Indels:	9
DB:	6	Gaps:	1

US-10-089-278-6_COPY_1_45 (1-45) x AX573109 (1-9602)

Qy	1	MetAspCysLeuThrAsnLeuArgSerAlaGluGlyValAspGlnAlaSerIysile	20
Db	6335	ATGGACTGTCACCAACCTCCGATCCGCTGAGGGTAAAGTTGACCAGGCGGCAAAATC	6394
Qy	21	Leu-----IleLeuLeuValAlaTrpTrp	28
Db	6395	CTAAGATCTAATTCACCCACCAGTCAGGCTGCCTATCAGAAAGTGGTGG	6445

Search completed: August 25, 2005, 16:35:50
Job time : 2869 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 25, 2005, 15:30:58 ; Search time 393 Seconds
(without alignments)
677.833 Million cell updates/sec

Title: US-10-089-278-6_COPY_1_45

Perfect score: 229
Sequence: 1 MDCLTLNLSAEGKVQASKI.....AWMGFTTAENVSTARAQPA 45

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q/cn2_1/USPTO spoool_h/US10089278/runat_25082005_131342_29638/app query.fasta_1.199
-DB=N Geneseq -QFW=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10089278 @CN2_1_1_470 @runat_25082005_131342_29638 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	229	100.0	906	5 AAF61513	AAf61513 DNA encod
2	229	100.0	927	5 AAF61510	AAf61510 DNA encod
3	229	100.0	946	5 AAF61512	AAf61512 DNA encod
4	229	100.0	990	5 AAF61511	AAf61511 DNA encod
5	229	100.0	1030	5 AAF61509	AAf61509 DNA encod

6	200	87.3	4776	2	AAx77617
7	200	87.3	4776	2	AAx77614
8	191	83.4	220	2	AAT04587
9	191	83.4	220	3	AAZ51114
10	183	79.9	1704	2	AAQ86456
11	108.5	47.4	9602	6	ABQ81148
12	108.5	47.4	9602	6	ABQ81149
13	104	45.4	5082	6	ABQ81150
14	63	27.5	978	6	ABK75572
15	62.5	27.3	648	11	ADM45388
16	62.5	27.3	665	11	ADM45384
17	62	27.1	423	2	AAV29840
18	62	27.1	2566	5	ABA19833
19	61	26.6	1681	10	ADC30409
20	61	26.6	110000	6	ABA03041.24
21	60	26.2	433	5	ABV50990
22	60	26.2	441	6	ABN61805
23	60	26.2	2034	13	ABD32699
24	60	26.2	25310	13	ABD32698
25	60	26.2	40392	6	ABL64734
26	60	26.2	40392	6	ABL65377
27	60	26.2	42104	11	ACN44606
28	60	26.2	160482	11	ACN43914
29	59.5	26.0	1956	6	ABX78664
30	59.5	26.0	1956	8	ABX56054
31	59	25.8	387	2	AAQ27141
32	59	25.8	81656	12	ADQ97876
33	58.5	25.5	417	2	AAQ03362
34	58.5	25.5	417	2	AAQ78565
35	58.5	25.5	417	2	AAT93544
36	58.5	25.5	417	2	AAV03201
37	58.5	25.5	417	2	AAV06412
38	58.5	25.5	417	2	AAV55093
39	58.5	25.5	417	2	AAV10287
40	58.5	25.5	417	2	AAQ90559
41	58.5	25.5	417	3	AAZ87947
42	58.5	25.5	417	3	AAQ65486
43	58.5	25.5	417	8	ABX63867
44	58.5	25.5	417	8	ABX81394
45	58.5	25.5	417	10	AAZ59288

ALIGNMENTS

RESULT 1
AAf61513
ID AAF61513 standard; DNA; 906 BP.
XX
AC AAF61513;
XX
DT 11-SEP-2003 (revised)
DT 25-JUN-2001 (first entry)
XX
DE DNA encoding SNV-env leader/human 6C3-scfv fusion construct.
XX
KW T lymphocyte; antibody; single chain variable antibody; scFv; human;
KW cell-targeting vector; antiviral; cytostatic; immunostimulant; AIDS;
KW gene therapy; immunization; diagnosis; T cell-associated disease; SCID;
KW acquired immune deficiency syndrome; severe combined immune deficiency;
KW T cell lymphoma; fusion construct; ds.
XX
OS Homo sapiens.
OS Spleen necrosis virus.
OS Chimeric.
XX
FH Key
CDS Location/Qualifiers
1..135
/*tag= a
/product= "SNV-env leader peptide"
/note= "No stop codon given"
136..906
/*tag= b
/product= "7B4-scfv"

```

FT      /note= "no stop codon given"
FT      /partial
XX      DE19946142-A1.
XX      29-MAR-2001.
XX      27-SEP-1999; 99DE-01046142.
XX      27-SEP-1999; 99DE-01046142.
XX      (BUND ) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.
XX      Cichutek K, Engelstaedter M;
XX      WPI; 2001-246140/26.
XX      P-PSDB; AAB70844.
XX      Cell-targeting vector selective for T lymphocytes, useful in gene therapy
PT      of e.g. acquired immune deficiency syndrome, encodes a single-chain
PT      variable antibody fragment.
XX      Claim 1; Fig 5; 18pp; German.
XX      This invention describes a novel cell-targeting vector (A) containing a
CC      DNA sequence (I) encoding a single-chain variable antibody fragment
CC      (scFv). The products of the invention have antiviral, cytostatic and
CC      immunostimulant activity and can be used in gene therapy, immunization
CC      and diagnosis particularly of T cell-associated diseases, specifically
CC      acquired immune deficiency syndrome (AIDS), severe combined immune
CC      deficiency (SCID) or T cell lymphoma. (A) target T cells, independently
CC      of the CD4 receptor, with high selectivity, 4-5 fold selectivity over
CC      human B cells, and 1000 fold selectivity over other human cells. A vector
CC      designated 7A5 encodes a 329 amino acid single-chain variable antibody
CC      fragment, fully defined in the specification. It was used to transform
CC      D17 (canine osteosarcoma cells susceptible to spleen necrosis virus
CC      (SNV)) cells, C8166 (human T lymphocyte) cells, and HeLa (human cervical
CC      carcinoma) cells. After 48 hours, the cells were stained with X-gal to
CC      determine transformation. The viral titer (infectious units/ml) was over
CC      1 million for D17, 1 million for C8166 but less than 100 for HeLa,
CC      showing the high selectivity for human T cells. This sequence encodes the
CC      SNV-env leader/human K6-scFv fusion construct used in the construction of
CC      novel cell targeting vectors described in the invention. (Updated on
CC      11-SEP-2003 to standardise OS field)
XX      Sequence 906 BP; 204 A; 246 C; 262 G; 194 T; 0 U; 0 Other;
SQ      Alignment Scores:
      Pred. No.:      8.18e-24      Length:      906
      Score:          229.00      Matches:      45
      Percent Similarity: 100.00%      Conservative: 0
      Best Local Similarity: 100.00%      Mismatches: 0
      Query Match:      100.00%      Indels:      0
      DB:              5      Gaps:      0
US-10-089-278-6_COPY_1_45 (1-45) x AAF61513 (1-906)
QY      1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyLysValaspGlnAlaSerLysIle 20
Db      1 ATGGAGCTGTCTACCAACCTCCGATCCGCTGAGGGTAAAGTTGACCAGCGGACAAATC 60
QY      21 LeuIleLeuValalaTrpTrpGlyPheGlyThrAlaGluValSerThrAlaArg 40
Db      61 CTAATTCTCTTGTGCTTGGTGGGGGTTTGGACCACTGCCGAAAGTTTCGACTGCCCGA 120
QY      41 AlaAlaGlnProAla 45
Db      121 GCGGCCCGACCGGCC 135
RESULT 2
AAF61510
ID      AAF61510 standard; DNA; 927 BP.
XX

```

```

AC      AAP61510;
XX      11-SEP-2003      (revised)
DT      25-JUN-2001      (first entry)
XX      DNA encoding SNV-env leader/human K6-scFv fusion construct.
XX      T lymphocyte; antibody; single chain variable antibody; scFv; human;
KW      cell-targeting vector; antiviral; cytostatic; immunostimulant; AIDS;
KW      gene therapy; immunization; diagnosis; T cell-associated disease; SCID;
KW      acquired immune deficiency syndrome; severe combined immune deficiency;
KW      T cell lymphoma; fusion construct; ds.
XX      Homo sapiens.
OS      Spleen necrosis virus.
OS      Chimeric.
XX      Key      Location/Qualifiers
FH      CDS      1..135
FT      /*tag= a
FT      /product= "SNV-env leader peptide"
FT      /note= "No stop codon given"
FT      136..927
FT      /*tag= b
FT      /product= "K6-scFv"
FT      /note= "no stop codon given"
FT      /partial
XX      DE19946142-A1.
XX      29-MAR-2001.
XX      27-SEP-1999; 99DE-01046142.
XX      27-SEP-1999; 99DE-01046142.
XX      (BUND ) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.
XX      Cichutek K, Engelstaedter M;
XX      WPI; 2001-246140/26.
XX      P-PSDB; AAB70841.
XX      Cell-targeting vector selective for T lymphocytes, useful in gene therapy
PT      of e.g. acquired immune deficiency syndrome, encodes a single-chain
PT      variable antibody fragment.
XX      Claim 1; Fig 2; 18pp; German.
XX      This invention describes a novel cell-targeting vector (A) containing a
CC      DNA sequence (I) encoding a single-chain variable antibody fragment
CC      (scFv). The products of the invention have antiviral, cytostatic and
CC      immunostimulant activity and can be used in gene therapy, immunization
CC      and diagnosis particularly of T cell-associated diseases, specifically
CC      acquired immune deficiency syndrome (AIDS), severe combined immune
CC      deficiency (SCID) or T cell lymphoma. (A) target T cells, independently
CC      of the CD4 receptor, with high selectivity, 4-5 fold selectivity over
CC      human B cells, and 1000 fold selectivity over other human cells. A vector
CC      designated 7A5 encodes a 329 amino acid single-chain variable antibody
CC      fragment, fully defined in the specification. It was used to transform
CC      D17 (canine osteosarcoma cells susceptible to spleen necrosis virus
CC      (SNV)) cells, C8166 (human T lymphocyte) cells, and HeLa (human cervical
CC      carcinoma) cells. After 48 hours, the cells were stained with X-gal to
CC      determine transformation. The viral titer (infectious units/ml) was over
CC      1 million for D17, 1 million for C8166 but less than 100 for HeLa,
CC      showing the high selectivity for human T cells. This sequence encodes the
CC      SNV-env leader/human K6-scFv fusion construct used in the construction of
CC      novel cell targeting vectors described in the invention. (Updated on
CC      11-SEP-2003 to standardise OS field)
XX      Sequence 927 BP; 205 A; 238 C; 277 G; 207 T; 0 U; 0 Other;
SQ      Alignment Scores:

```

```

Pred. No.: 8.42e-24 Length: 927
Score: 229.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x AAF61510 (1-927)

QY 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysile 20
DB 1 ATGGACTGCTCACCACCTCCGTCGCTGAGGGTAAGTTGACCGAGCGGAGCAAAATC 60

QY 21 LeuileLeuValAlaTrpTrpGlyPheGlyThrThrAlaGluValSerThrAlaArg 40
DB 61 CTAATTCTCTTGCTGGCTGGTGGGGTTTGGACCACTGCCGAAGTTTCGACTGCCCGA 120

QY 41 AlaAlaGlnProAla 45
DB 121 GCGGCCCGAGCGGCC 135

RESULT 3
AAF61512
ID AAF61512 standard; DNA; 946 BP.
XX
AC AAF61512;
XX
DT 11-SEP-2003 (revised)
DT 25-JUN-2001 (first entry)
XX
DE DNA encoding SNV-env leader/human 7E4-scFv fusion construct.
XX
KW T lymphocyte; antibody; single chain variable antibody; scFv; human;
KW cell-targeting vector; antiviral; cytostatic; immunostimulant; AIDS;
KW gene therapy; immunization; diagnosis; T cell-associated disease; SCID;
KW acquired immune deficiency syndrome; severe combined immune deficiency;
KW T cell lymphoma; fusion construct; ds.
XX
OS Homo sapiens.
OS Spleen necrosis virus.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..135
FT /tag= a
FT /product= "SNV-env leader peptide"
FT /note= "No stop codon given"
FT 136..946
FT /tag= b
FT /product= "7E4-scFv"
FT /note= "no stop codon given"
FT /partial
XX
DE19946142-A1.
XX
XX
PD 29-MAR-2001.
XX
XX
PF 27-SEP-1999; 99DE-01046142.
XX
XX
PR 27-SEP-1999; 99DE-01046142.
XX
XX
PA (BUND ) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.
XX
XX
PI Cichutek K, Engeltaedter M;
XX
DR WPI; 2001-246140/26.
DR P-PSDB; AAB70843.
XX
XX
PT Cell-targeting vector selective for T lymphocytes, useful in gene therapy
PT of e.g. acquired immune deficiency syndrome, encodes a single-chain
PT variable antibody fragment.
XX
PS Claim 1; Fig 4; 18pp; German.

```

```

XX This invention describes a novel cell-targeting vector (A) containing a
CC DNA sequence (1) encoding a single-chain variable antibody fragment
CC (scFv). The products of the invention have antiviral, cytostatic and
CC immunostimulant activity and can be used in gene therapy, immunization
CC and diagnosis particularly of T cell-associated diseases, specifically
CC acquired immune deficiency syndrome (AIDS), severe combined immune
CC deficiency (SCID) or T cell lymphoma. (A) target T cells, independently
CC of the CD4 receptor, with high selectivity, 4-5 fold selectivity over
CC human B cells, and 1000 fold selectivity over other human cells. A vector
CC designated 7A5 encodes a 329 amino acid single-chain variable antibody
CC fragment, fully defined in the specification. It was used to transform
CC Di7 (canine osteosarcoma cells susceptible to spleen necrosis virus
CC (SNV)) cells, C8166 (human T lymphocyte) cells, and HeLa (human cervical
CC carcinoma) cells. After 48 hours, the cells were stained with X-gal to
CC determine transfection. The viral titer (infectious units/ml) was over
CC 1 million for Di7, 1 million for C8166 but less than 100 for HeLa,
CC showing the high selectivity for human T cells. This sequence encodes the
CC SNV-env leader/human 7E4-scFv fusion construct used in the construction
CC of novel cell targeting vectors described in the invention. (Updated on
CC 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 946 BP; 199 A; 249 C; 276 G; 222 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 8.64e-24 Length: 946
Score: 229.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x AAF61512 (1-946)

QY 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysile 20
DB 1 ATGGACTGCTCACCACCTCCGTCGCTGAGGGTAAGTTGACCGAGCGGAGCAAAATC 60

QY 21 LeuileLeuValAlaTrpTrpGlyPheGlyThrThrAlaGluValSerThrAlaArg 40
DB 61 CTAATTCTCTTGCTGGCTGGTGGGGTTTGGACCACTGCCGAAGTTTCGACTGCCCGA 120

QY 41 AlaAlaGlnProAla 45
DB 121 GCGGCCCGAGCGGCC 135

RESULT 4
AAF61511
ID AAF61511 standard; DNA; 990 BP.
XX
AC AAF61511;
XX
DT 11-SEP-2003 (revised)
DT 25-JUN-2001 (first entry)
XX
DE DNA encoding SNV-env leader/human 7B2-scFv fusion construct.
XX
KW T lymphocyte; antibody; single chain variable antibody; scFv; human;
KW cell-targeting vector; antiviral; cytostatic; immunostimulant; AIDS;
KW gene therapy; immunization; diagnosis; T cell-associated disease; SCID;
KW acquired immune deficiency syndrome; severe combined immune deficiency;
KW T cell lymphoma; fusion construct; ds.
XX
OS Homo sapiens.
OS Spleen necrosis virus.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..135
FT /tag= a
FT /product= "SNV-env leader peptide"
FT /note= "No stop codon given"
FT 136..990
FT /tag= b
FT /product= "7B2-scFv"
FT /note= "no stop codon given"
FT /partial
XX
DE19946142-A1.
XX
XX
PD 29-MAR-2001.
XX
XX
PF 27-SEP-1999; 99DE-01046142.
XX
XX
PR 27-SEP-1999; 99DE-01046142.
XX
XX
PA (BUND ) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.
XX
XX
PI Cichutek K, Engeltaedter M;
XX
DR WPI; 2001-246140/26.
DR P-PSDB; AAB70843.
XX
XX
PT Cell-targeting vector selective for T lymphocytes, useful in gene therapy
PT of e.g. acquired immune deficiency syndrome, encodes a single-chain
PT variable antibody fragment.
XX
PS Claim 1; Fig 4; 18pp; German.

```

```

FT      /*tag= b
FT      /product= "7B2-scFv"
FT      /note= "no stop codon given"
FT      /partial
PN      DE19946142-A1.
XX      29-MAR-2001.
XX
XX      27-SEP-1999; 99DE-01046142.
XX
XX      27-SEP-1999; 99DE-01046142.
XX
XX      (BUND ) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.
XX
XX      Cichutek K, Engelstaedter M;
XX
XX      WPI: 2001-246140/26.
XX      P-PSDB; AAB70842.
XX
XX      Cell-targeting vector selective for T lymphocytes, useful in gene therapy
PT      of e.g. acquired immune deficiency syndrome, encodes a single-chain
PT      variable antibody fragment.
XX
XX      Claim 1; Fig 3; 18pp; German.
XX
XX      This invention describes a novel cell-targeting vector (A) containing a
CC      DNA sequence (I) encoding a single-chain variable antibody fragment
CC      (scFv). The products of the invention have antiviral, cytostatic and
CC      immunostimulant activity and can be used in gene therapy, immunization
CC      and diagnosis particularly of T cell-associated diseases, specifically
CC      acquired immune deficiency syndrome (AIDS), severe combined immune
CC      deficiency (SCID) or T cell lymphoma. (A) target T cells, independently
CC      of the CD4 receptor, with high selectivity, 4-5 fold selectivity over
CC      human B cells, and 1000 fold selectivity over other human cells. A vector
CC      designated 7A5 encodes a 329 amino acid single-chain variable antibody
CC      fragment, fully defined in the specification. It was used to transform
CC      D17 (canine osteosarcoma cells susceptible to spleen necrosis virus
CC      (SNV)) cells, C8166 (human T lymphocyte) cells, and Hela (human cervical
CC      carcinoma) cells. After 48 hours, the cells were stained with X-gal to
CC      determine transformation. The viral titer (infectious units/ml) was over
CC      1 million for D17, 1 million for C8166 but less than 100 for Hela,
CC      showing the high selectivity for human T cells. This sequence encodes the
CC      SNV-env leader/human 7B2-scFv fusion construct used in the construction
CC      of novel cell targeting vectors described in the invention. (Updated on
CC      11-SEP-2003 to standardise OS field)
XX
XX      Sequence 990 BP; 210 A; 263 C; 289 G; 228 T; 0 U; 0 Other;
SQ

```

```

Alignment Scores:
Pred. No.: 9.15e-24 Length: 990
Score: 229.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

```

US-10-089-278-6_COPY_1_45 (1-45) x AAF61501 (1-990)

```

QY      1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysile 20
DB      1 ATGACCTGTCTCACCACCTCCGATCCGTCGAGGTAAAGTTGACCGCGGACGACAAATC 60
QY      21 LeuIleLeuValAlaTrpTrpGlyPheGlyThrAlaGluValSerThrAlaArg 40
DB      61 CTAATTCCTCTGGCTTGGTGGGGTGGTGGGACCACTGCCAGTTCGACTGCCCGA 120
QY      41 AlaAlaGlnProAla 45
DB      121 GCGGCCGACCGCGCC 135

```

RESULT 5
AAF61509

```

ID      AAF61509 standard; DNA; 1030 BP.
XX
XX      AAF61509;
XX
XX      11-SEP-2003 (revised)
XX      25-JUN-2001 (first entry)
XX
XX      DNA encoding SNV-env leader/human 7A5-scFv fusion construct.
XX
XX      T lymphocyte; antibody; single chain variable antibody; scFv; human;
KW      cell-targeting vector; antiviral; cytostatic; immunostimulant; AIDS;
KW      gene therapy; immunization; diagnosis; T cell-associated disease; SCID;
KW      acquired immune deficiency syndrome; severe combined immune deficiency;
KW      T cell lymphoma; fusion construct; ds.
XX
XX      Homo sapiens.
OS      Spleen necrosis virus.
OS      Chimeric.
XX
XX      Key Location/Qualifiers
XX      CDS 44..178
XX      /tag= a
XX      /product= "SNV-env leader peptide"
XX      /note= "No stop codon given"
XX      179..1030
XX      /tag= b
XX      /product= "7A5-scFv"
XX      /note= "no stop codon given"
XX      /partial
XX
XX      DE19946142-A1.
XX
XX      29-MAR-2001.
XX
XX      27-SEP-1999; 99DE-01046142.
XX
XX      27-SEP-1999; 99DE-01046142.
XX
XX      (BUND ) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.
XX
XX      Cichutek K, Engelstaedter M;
XX
XX      WPI: 2001-246140/26.
XX      P-PSDB; AAB70840.
XX
XX      Cell-targeting vector selective for T lymphocytes, useful in gene therapy
PT      of e.g. acquired immune deficiency syndrome, encodes a single-chain
PT      variable antibody fragment.
XX
XX      Claim 1; Fig 1; 18pp; German.
XX
XX      This invention describes a novel cell-targeting vector (A) containing a
CC      DNA sequence (I) encoding a single-chain variable antibody fragment
CC      (scFv). The products of the invention have antiviral, cytostatic and
CC      immunostimulant activity and can be used in gene therapy, immunization
CC      and diagnosis particularly of T cell-associated diseases, specifically
CC      acquired immune deficiency syndrome (AIDS), severe combined immune
CC      deficiency (SCID) or T cell lymphoma. (A) target T cells, independently
CC      of the CD4 receptor, with high selectivity, 4-5 fold selectivity over
CC      human B cells, and 1000 fold selectivity over other human cells. A vector
CC      designated 7A5 encodes a 329 amino acid single-chain variable antibody
CC      fragment, fully defined in the specification. It was used to transform
CC      D17 (canine osteosarcoma cells susceptible to spleen necrosis virus
CC      (SNV)) cells, C8166 (human T lymphocyte) cells, and Hela (human cervical
CC      carcinoma) cells. After 48 hours, the cells were stained with X-gal to
CC      determine transformation. The viral titer (infectious units/ml) was over
CC      1 million for D17, 1 million for C8166 but less than 100 for Hela,
CC      showing the high selectivity for human T cells. This sequence encodes the
CC      SNV-env leader/human 7A5-scFv fusion construct used in the construction
CC      of novel cell targeting vectors described in the invention. (Updated on
CC      11-SEP-2003 to standardise OS field)
XX
XX      Sequence 1030 BP; 232 A; 244 C; 304 G; 250 T; 0 U; 0 Other;
SQ

```



```
Alignment Scores:
Pred. No.:      9.62e-24      Length:      1030
Score:          229.00      Matches:      45
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              5          Gaps:      0

US-10-089-278-6_COPY_1_45 (1-45) x AAF61509 (1-1030)

QY      1 MetAspCysLeuThrLeuArgSerAlaGluGlyValAspGlnAlaSerIlySile 20
      |||||
Db      44 ATGACTGTCCTACCAACCTCCGATCCGCTGAGGGTAAAGTTGACCAGCGAGCAAAATC 103
      |||||

QY      21 LeuIleLeuValAlaIleTrpGlyPheGlyThrThrAlaGluValSerThrAlaArg 40
      |||||
Db      104 CTAATTCCTCTGCTGGCTGGTGGGGGTTGGGACCACTGGCCGAGTTTCGACTGCCCGA 163
      |||||

QY      41 AlaAlaGlnProAla 45
      |||||
Db      164 GCGGCCCGAGCGGCC 178
      |||||

RESULT 6
AAX77617
ID      AAX77617 standard; DNA; 4776 BP.
XX
AC      AAX77617;
XX
DT      13-AUG-1999 (first entry)
XX
DE      Expression construct pTC53 DNA.
XX
KW      Pseudo-type retroviral vector; surface capsid protein; virus core;
KW      retroviral packaging cell; psi-negative expression construct; gag gene;
KW      pol gene; cell-specific transduction; cell targeting; gene therapy;
KW      vaccination; diagnosis; cystic fibrosis; ADA-deficiency; HIV-1 infection;
KW      chronic granulomatosis; ss.
XX
OS      Spleen necrosis virus.
OS      Mus sp.
OS      Synthetic.
XX

FH      Key
CDS      Location/Qualifiers
          64..102
          /tag= a
          /note= "CDS 1"
          123..217
          /tag= b
          /note= "CDS 2"
          251..292
          /tag= c
          /note= "CDS 3"
          326..364
          /tag= d
          /note= "CDS 4"
          376..408
          /tag= e
          /note= "CDS 5"
          829..876
          /tag= f
          /note= "CDS 6"
          883..957
          /tag= g
          /note= "CDS 7"
          1025..1123
          /tag= h
          /note= "CDS 8"
          1759..1803
          /tag= i
          /note= "CDS 9"
          1840..1887
          /tag= j

FT      CDS
FT      /note= "CDS 10"
FT      1911..1961
FT      /tag= k
FT      /note= "CDS 11"
FT      1967..1981
FT      /tag= l
FT      /note= "CDS 12"
FT      1990..2010
FT      /tag= m
FT      /note= "CDS 13"
FT      2061..2099
FT      /tag= n
FT      /note= "CDS 14"
FT      2157..2237
FT      /tag= o
FT      /note= "CDS 15"
FT      2253..2402
FT      /tag= p
FT      /note= "CDS 16"
FT      2757..2771
FT      /tag= q
FT      /note= "CDS 17"
FT      3223..3252
FT      /tag= r
FT      /note= "CDS 18"
FT      3299..3334
FT      /tag= s
FT      /note= "CDS 19"
FT      3451..3717
FT      /tag= t
FT      /note= "CDS 20"
FT      3724..3894
FT      /tag= u
FT      /note= "CDS 21"
FT      3904..4053
FT      /tag= v
FT      /note= "CDS 22"
FT      4133..4153
FT      /tag= w
FT      /note= "CDS 23"
FT      4167..4250
FT      /tag= x
FT      /note= "CDS 24"
FT      4251..4268
FT      /tag= y
FT      /note= "CDS 25"
FT      4336..4416
FT      /tag= z
FT      /note= "CDS 26"
FT      4420..4455
FT      /tag= aa
FT      /note= "CDS 27"
FT      4460..4555
FT      /tag= ab
FT      /note= "CDS 28"
FT      4567..4776
FT      /tag= ac
FT      /note= "CDS 29"
XX      WO9928488-A2.
XX
XX      10-JUN-1999.
XX
XX      27-NOV-1998; 98WO-DE003542.
XX
XX      28-NOV-1997; 97DE-01052855.
XX
XX      (BUND ) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.
XX      Cichutek K, Merget-Millitzer H;
XX      WPI; 1999-358132/30.
XX
```

PT Pseudo-type retroviral vectors with modified surface capsid proteins.
XX Disclosure; Fig 4A-B; 4lpp; German.
XX
CC This invention describes novel pseudo-type retroviral vectors with
CC modified surface capsid proteins. The vectors of the invention consist
CC essentially of a virus core chosen from the group of murine leukemia
CC virus (MLV), human immunodeficiency virus (HIV), simian immunodeficiency
CC virus (SIV), lentivirus or foamyvirus and a virus capsid protein from
CC spleen necrotic virus (SNV). The invention also describes a retroviral
CC packaging cell for the retroviral vector above, and also transformed with
CC one or more psi-negative expression constructs, the gag and pol gene
CC products of MLV, HIV, SIV or foamyvirus, or also with a psi-negative SNV-
CC env expression construct and/or psi-negative SNV-ENV foreign polypeptide-
CC SNV-HIV-ENV or SNV-SIV-ENV expression construct. The pseudo-type
CC retroviral vectors with modified surface capsid proteins are suitable for
CC cell-specific transduction of a selected mammal cell type (cell
CC targeting). The methods are useful for the production of the pseudo-type
CC retroviral vectors and for gene transfer in selected cell types. The
CC vectors can be used in medicaments for gene therapy, vaccination or
CC diagnosis. They are particularly useful for therapy of cystic fibrosis,
CC ADA-deficiency, chronic granulomatosis or HIV-1 infection. This sequence
CC represents the expression construct pTC53 which is composed from the SNV
CC ENV protein and a murine derived scFv fragment. This sequence encodes the
CC protein fragments represented in AAY08848-Y08877
XX
SQ Sequence 4776 BP; 1199 A; 1178 C; 1245 G; 1154 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1-22e-18 Length: 4776
Score: 200.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 87.34% Indels: 0
DB: 2 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x AAX77617 (1-4776)

QY 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIle 20
Db 1025 ATGGACTGTCTCACCACCTCCGATCCGCTGAGGGTAAAGTTGACCAGCGGACAAATC 1084
QY 21 LeuIleLeuValAlaTrpTrpGlyPheGlyThrAlaGluValSerThrAla 39
Db 1085 CTAATTCCTCTGTGGCTTGGUGGGGTTGGGACCACCTGCCGAGTTTCGACTGCC 1141

RESULT 7

AAX77614
ID AAX77614 standard; DNA; 4776 BP.
XX
XX AAX77614;
XX
DT 13-AUG-1999 (first entry)
XX
DE Expression construct pTC53 DNA.
XX
KW Cell-specific retroviral vector; antibody domain; vaccination; scFv;
KW cell-specific transduction; B cell RNA; variable region; heavy chain;
KW light chain; immunoglobulin; psi-negative; retroviral Env protein;
KW capsid protein; cell targeting; gene therapy; diagnosis; cystic fibrosis;
KW ADA-deficiency; chronic granulomatosis; HIV-1 infection; da.
XX
OS Spleen necrosis virus.
OS Mus sp.
OS Synthetic.
XX
XX Location/Qualifiers
FH Key 64..102
FT CDS /tag= a
FT /note= "CDS 1"
FT 123..217
FT CDS /tag= b
FT /note= "CDS 2"
FT

CDS FT 251..292
FT /tag= c
FT /note= "CDS 3"
CDS FT 326..364
FT /tag= d
FT /note= "CDS 4"
CDS FT 376..408
FT /tag= e
FT /note= "CDS 5"
CDS FT 829..876
FT /tag= f
FT /note= "CDS 6"
CDS FT 883..957
FT /tag= g
FT /note= "CDS 7"
CDS FT 1025..1723
FT /tag= h
FT /note= "CDS 8"
CDS FT 1759..1803
FT /tag= i
FT /note= "CDS 9"
CDS FT 1840..1887
FT /tag= j
FT /note= "CDS 10"
CDS FT 1911..1961
FT /tag= k
FT /note= "CDS 11"
CDS FT 1967..1981
FT /tag= l
FT /note= "CDS 12"
CDS FT 1990..2010
FT /tag= m
FT /note= "CDS 13"
CDS FT 2061..2099
FT /tag= n
FT /note= "CDS 14"
CDS FT 2157..2237
FT /tag= o
FT /note= "CDS 15"
CDS FT 2253..2402
FT /tag= p
FT /note= "CDS 16"
CDS FT 2757..2771
FT /tag= q
FT /note= "CDS 17"
CDS FT 3223..3252
FT /tag= r
FT /note= "CDS 18"
CDS FT 3299..3334
FT /tag= s
FT /note= "CDS 19"
CDS FT 3451..3717
FT /tag= t
FT /note= "CDS 20"
CDS FT 3724..3894
FT /tag= u
FT /note= "CDS 21"
CDS FT 3904..4053
FT /tag= v
FT /note= "CDS 22"
CDS FT 4133..4153
FT /tag= w
FT /note= "CDS 23"
CDS FT 4167..4250
FT /tag= x
FT /note= "CDS 24"
CDS FT 4251..4268
FT /tag= y
FT /note= "CDS 25"
CDS FT 4336..4416
FT /tag= z
FT /note= "CDS 26"
CDS FT 4420..4455

FT /*tag= aa
 FT /note= "CDS 27"
 FT 4460. .4555
 FT /*tag= ab
 FT /note= "CDS 28"
 FT 4567. .4776
 FT /*tag= ac
 FT /note= "CDS 29"
 XX
 PN WO9928489-A2.
 XX
 XX 10-JUN-1999.
 PD
 XX
 XX 27-NOV-1998; 98WO-DE003543.
 PF
 XX 28-NOV-1997; 97DE-01052854.
 PR
 XX (BUND) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.
 PA
 XX Cichutek K, Engelstaedter M;
 PI
 XX WPI; 1999-371131/31.
 DR
 XX
 XX Cell-specific retroviral vectors with antibody domains suitable for cell-
 PT specific transduction of selected mammal cell types - useful for
 PT vaccination and gene therapy for treatment of, e.g. cystic fibrosis.
 PT
 XX Disclosure; Fig 4A-B; 38pp; German.
 PS
 XX This invention describes the construction of novel cell-specific
 CC retroviral vectors with antibody domains suitable for cell-specific
 CC transduction of selected mammal cell types. The invention describes a
 CC method to produce cell-specific retroviral vectors which consists
 CC essentially of the following steps: (a) immunization of a mammal with one
 CC or more cell populations (b) isolation of RNA from the immunized mammal,
 CC especially the B cell RNA (c) production of a cDNA strand of the variable
 CC region of the heavy and light chains of the immunoglobulins isolated from
 CC the RNA by RT-PCR with primers for the respective immunoglobulin chains,
 CC where the primer nucleic acid sequences are for an oligopeptide linker
 CC (d) ligation of the cDNA strain to scFV-cDNA (e) ligation of the scFV
 CC cDNA in a phagemid vector and transformation of a host bacterium with the
 CC vector (f) isolation of phage, by selection of phage that bind to the
 CC cell population used in step (a) (g) cleavage of the scFV coding DNA
 CC fragments from the cell-specific phage and ligation into a psi-negative
 CC retroviral Env-expression vector (h) transformation of a Env-ScFV
 CC expression vector to be maintained in a packaging cell and (i) isolation
 CC of a packaging cell with the retroviral vectors. The pseudo-type
 CC retroviral vectors with modified surface capsid proteins are suitable for
 CC cell-specific transduction of a selected mammal cell type (cell
 CC targeting). The methods are useful for the production of the pseudo-type
 CC retroviral vectors and for gene transfer in selected cell types. The
 CC vectors can be used in medicaments for gene therapy, vaccination or
 CC diagnosis. They are particularly useful for therapy of cystic fibrosis,
 CC ADA-deficiency, chronic granulomatosis or HIV-1 infection. This sequence
 CC represents the expression construct pUC53 which is composed from the SNV
 CC ENV protein and a murine derived scFV fragment. This sequence encodes the
 CC protein fragments represented in AAY08761-Y08790
 XX
 SQ Sequence 4776 BP; 1199 A; 1178 C; 1245 G; 1154 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,22e-18 Length: 4776
 Score: 200.00 Matches: 39
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 87.34% Indels: 0
 DB: 2 Gaps: 0
 US-10-089-278-6_COPY_1_45 (1-45) x AAY77614 (1-4776)
 Qy 1 MetAspCysLeuThrAnLeuAvrSerAlaGluGlyValAspGlnAlaSerIlylle 20
 Db 1025 ATGGACTGTCTCCACCACTCCGATCCGCTGAGGGTAAAGTTGACCGAGCGCAAAATC 1084

QY 21 LeuileLeuLeuValalaTrpTrpGlyPheGlyThrAlaGluValSerThrAla 39
 Db 1085 CTAATTCCTCTTGTGGCTTGGTGGGGTTTGGACCACTGCCGAAGTTTCGACTGCC 1141
 RESULT 8
 AAT04587
 ID AAT04587 standard; DNA; 220 BP.
 XX
 AC AAT04587;
 XX
 DT 09-APR-1996 (first entry)
 XX
 DE Spleen necrosis virus leader sequence.
 XX
 KW Antibody; scFv; targeting peptide; retroviral vector; gene therapy;
 KW adenosine deaminase deficiency; ADA; cancer; haptan dinitrophenol; B6.2;
 KW HeLa; Col-1; spleen necrosis virus; SNV; ss.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 85..207
 FT /*tag= a
 FT /product= "MLV-pro"
 XX
 PN WO9523846-A1.
 XX
 XX 08-SEP-1995.
 XX
 PF 03-MAR-1995; 95WO-US002537.
 XX
 PR 04-MAR-1994; 94US-00205980.
 XX
 PA (UYNE-) UNIV NEW JERSEY.
 XX
 PI Dornburg RC;
 XX
 DR WPI; 1995-320563/41.
 DR P-PSDB; AAR80070.
 XX
 XX Retroviral vectors for use in cell specific gene transfer - contain
 FT antibody-envelope and wild-type envelope-fusion proteins.
 XX
 PS Example; Fig 5; 35pp; English.
 XX
 CC This sequence represents the coding sequence for the leader sequence of
 CC the spleen necrosis virus (SNV). This sequence is included in an expression
 CC expression vector pPD114. This sequence is included in an expression
 CC vector (pTC13) and is linked to the B6.2 sequence amplified by AAT04585
 CC and AAT04586. B6.2 acts as a targeting peptide (TP), and will direct the
 CC retroviral vector to a cell-surface protein expressed on the surface of
 CC various human cancers (e.g. HeLa and Col-1). By using different TP's,
 CC that recognise different cell surface antigens, the retroviral vectors
 CC containing these TP's can be used in a cell type specific method for
 CC introducing genes into cells. These retroviral vectors can be used in the
 CC gene therapy of human genetic diseases including, adenosine deaminase
 CC (ADA) deficiency, and in clinical trials to cure cancer. A wild type
 CC envelope can be used in addition to the altered vector, and will act as a
 CC helper molecule. The helper function enhances the infection of cells by
 CC the retroviral vector
 XX
 SQ Sequence 220 BP; 55 A; 61 C; 60 G; 44 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5.18e-19 Length: 220
 Score: 191.00 Matches: 37
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 83.41% Indels: 0
 DB: 2 Gaps: 0
 US-10-089-278-6_COPY_1_45 (1-45) x AAT04587 (1-220)

QY 1 MetAseCysLeuThrAsnLeuArgSerAlaGluGlyValaspGlnAlaSerLysIle 20
 DB 85 ATGGACTGCTCACCACCTCCGATCCGCTGAGGGTAAAGTTGACACGCGAGCAAAATC 144
 QY 21 LeuIleLeuValAlaTrpTrpGlyPheGlyThrAlaGluValSer 37
 DB 145 CTAATTCCTCTGTGGCTTGGTGGGGGTTGGGACCACTGCCGAAGTTTCG 195

RESULT 9
 AAZ51114
 ID AAZ51114 standard; DNA; 220 BP.
 AC AAZ51114;
 XX
 DT 05-JUN-2000 (first entry)
 DE Eucaryotic gene expression vector pTC13.
 XX
 KW Retroviral vector; Spleen Necrosis Virus; SNV;
 KW antibody-envelope fusion protein; retroviral envelope protein;
 KW gene therapy; antigen binding site; single chain antibody; scFv;
 KW dinitrophenol; DNP; eucaryotic gene expression vector; pTC13; B6.2 gene;
 KW tumour cell; cell-surface protein; cell specific gene transfer; ds.
 XX
 OS Murine leukemia virus.
 OS Mastadenovirus.
 OS Spleen necrosis virus.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT promoter 1..84
 FT /*tag= a
 FT /label= Murine_leukemia_virus_promoter/enhancer
 FT CDS 85..207
 FT /*tag= b
 FT /product= "Protein encoded by pTC13 vector"
 FT sig_peptide 187..220
 FT /*tag= c
 FT /label= SNV_leader_sequence
 XX
 PN WO200009730-A2.
 XX
 PD 24-FEB-2000.
 XX
 PF 10-AUG-1999; 99WO-US018141.
 XX
 PR 17-AUG-1998; 98US-00135121.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 PI Dornburg RC;
 XX
 XX WPI: 2000-224358/19.
 DR P-PSDB; AAY70110.
 XX
 XX Cell specific gene transfer using retroviral vectors containing antibody-
 FT envelope fusion proteins and wild type envelope proteins.
 XX
 XX Example; Fig 5; 45pp; English.
 PS
 CC The patent discloses a novel retroviral vector, particularly Spleen
 CC Necrosis Virus (SNV) vector, having target cell specificity. The vector
 CC has a targeting envelope which is a chimeric protein consisting of an
 CC antigen binding site of an antibody (e.g. anti-DNP-scFv) or another
 CC peptide that binds to a specific cell surface protein, fused to the
 CC carboxy terminal part of the retroviral envelope protein. The presence of
 CC the wild type envelope protein serves as a helper molecule to improve or
 CC supplement a functional membrane fusion domain. The antigen binding site
 CC replaces the natural viral receptor binding site. The retroviral vector
 CC is used for cell specific gene transfer, especially in gene therapy. The
 CC invention overcomes the restricted host range limitation of retroviral
 CC vectors. The present sequence is a eucaryotic gene expression vector

CC pTC13 derived from another vector pRD114. The vector contains a gene
 CC fragment encoding an endoplasmic reticulum (ER) recognition signal
 CC sequence, a murine leukaemia virus promoter/enhancer sequence, an
 CC adenovirus tripartite leader sequence, an SNV leader sequence, and a SV40
 CC poly A signal sequence. The vector is used to construct a targeting
 CC envelope directed to a cell-surface protein expressed on several human
 CC tumour cells. The targeting envelope contains a single chain antibody
 CC B6.2 gene fused to SNV envelope gene
 XX
 SQ Sequence 220 BP; 55 A; 61 C; 60 G; 44 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5,18e-19 Length: 220
 Score: 191.00 Matches: 37
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 83.41% Indels: 0
 DB: 3 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x AAZ51114 (1-220)
 QY 1 MetAseCysLeuThrAsnLeuArgSerAlaGluGlyValaspGlnAlaSerLysIle 20
 DB 85 ATGGACTGCTCACCACCTCCGATCCGCTGAGGGTAAAGTTGACACGCGAGCAAAATC 144
 QY 21 LeuIleLeuValAlaTrpTrpGlyPheGlyThrAlaGluValSer 37
 DB 145 CTAATTCCTCTGTGGCTTGGTGGGGGTTGGGACCACTGCCGAAGTTTCG 195

RESULT 10
 AAQ86456
 ID AAQ86456 standard; DNA; 1704 BP.
 XX
 AC AAQ86456;
 XX
 DT 29-SEP-1995 (first entry)
 XX
 DE Spleen necrosis virus env gene.
 XX
 KW Avipoxvirus; poxvirus; reticuloendotheliosis virus; REV; vaccine;
 KW poultry; ds.
 XX
 OS Spleen necrosis virus.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1704
 FT /*tag= a
 XX
 PN US5403582-A.
 XX
 PD 04-APR-1995.
 XX
 PF 21-JAN-1993; 93US-00007282.
 XX
 PR 21-JAN-1993; 93US-00007282.
 XX
 PA (JAPG) NIPPON ZEON KK.
 PA (USDA) US SEC OF AGRIC.
 XX
 PI Yanagida N, Nazerian K, Calvert JG, Witter RL;
 XX
 DR WPI: 1995-146769/19.
 DR P-PSDB; AAR71700.
 XX
 XX New Avipoxvirus expressing glycoprotein of spleen necrosis virus - used
 PT in vaccines to protect poultry against avian reticuloendotheliosis
 PT retrovirus related diseases.
 XX
 PS Claim 3; Col 17-20; 19pp; English.
 XX
 CC A claimed fowlpox virus expresses a gene (given in AAQ86456) encoding an
 CC envelope glycoprotein (AAR71700) of the spleen necrosis virus strain of
 CC REV, the gene being inserted at position 25 or 29 of the fowlpox virus

US-10-089-278-6 COPY 1 45 (1-45) X ABO81149 (1-9602)

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 45.41% Indels: 0
DB: 6 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x ABQ81150 (1-5082)

Qy 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysile 20
Db 5019 ATGGACTGTCTCACCACCTCCGATCCGCTGAGGGTAAAGTTGACCGAGCGGAGCAAAATC 5078
Qy 21 Leu 21
Db 5079 CTA 5081

RESULT 14
ABK75572
ID ABK75572 standard; DNA; 978 BP.
XX
AC ABK75572;
XX
DT 13-AUG-2002 (first entry)
XX
DE Bacillus licheniformis genomic sequence tag (GST) #2863.
XX
KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX
OS Bacillus licheniformis.
XX
PN WO200229113-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-US031437.
XX
PR 06-OCT-2000; 2000US-00680598.
XX
PA (NOVO) NOVOZYMES BIOTECH INC.
XX
PA (NOVO) NOVOZYMES AS.
XX
PI Berka R, Clausen IG;
XX
DR WPI; 2002-416684/44.
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
array.
XX
PS Claim 4; SEQ ID NO 2863; 200pp; English.
XX

The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive follow
CC -up characterisation is unnecessary, when one spot on an array equals one
CC gene or one open reading frame, since sequence information is available.
CC This sequence represents a genomic sequence tag (GST) used in the method
CC of the invention. Note: The sequence data for this patent did not form

CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 978 BP; 220 A; 205 C; 226 G; 327 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 21.7 Length: 978
Score: 63.00 Matches: 12
Percent Similarity: 68.00% Conservative: 5
Best Local Similarity: 48.00% Mismatches: 8
Query Match: 27.51% Indels: 0
DB: 6 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x ABK75572 (1-978)

Qy 13 LysValAspGlnAlaSerLysileLeuLeuValAlaTrpTrpGlyPheGlyThr 32
Db 811 AAAATGGATCGTCCGATCCAGATTCTCGTGATTTTATTTCGGCATCTGGGGATTGGGCACA 870
Qy 33 ThrAlaGluValSer 37
Db 871 GCCGCTGCTTCGAGT 885

RESULT 15
ADM45388/c
ID ADM45388 standard; DNA; 648 BP.
XX
AC ADM45388;
XX
DT 03-JUN-2004 (first entry)
XX
DE Insect resistance associated DNA sequence SeqID795.
XX
KW insect resistant phenotype; plant protectant; gene therapy;
KW Arabidopsis thaliana; Nicotiana benthamiana; Oryza sativa;
KW Papaver rhoeas; rice; insect resistance; insect-resistant plant; ds.
XX
OS Unidentified.
XX
PN WO2003020025-A2.
XX
PD 13-MAR-2003.
XX
PF 30-AUG-2002; 2002WO-US027882.
XX
PR 31-AUG-2001; 2001US-0316319P.
XX
PA (DOWC) DOW CHEM CO.
XX
PI Shukla V, Meade T, Larrinua I;
XX
DR WPI; 2003-290133/28.
XX

New isolated nucleic acid having expression that results in an insect
PT resistant phenotype, useful for conferring insect resistance and for
PT producing insect-resistant plants.
XX
PS Claim 1; SEQ ID NO 795; 396pp; English.
XX

This invention relates to a novel isolated nucleic acid comprising, or
CC hybridising under low stringent conditions to, any of the 1314 nucleic
CC acid sequences given in the specification, where the expression of the
CC nucleic acid in a plant results in an insect resistant phenotype. The
CC invention may be useful as a plant protectant or for gene therapy. The
CC genes are derived from Arabidopsis thaliana, Nicotiana benthamiana, Oryza
CC sativa and Papaver rhoeas. The isolated nucleic acid and vector are
CC useful for conferring insect resistance and for producing insect-
CC resistant plants. The present sequence is that of a DNA sequence of the
CC invention which may confer insect resistance to plants.
XX
SQ Sequence 648 BP; 146 A; 175 C; 175 G; 152 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	15.2	Length:	648
Score:	42.50	Matches:	17
Percent Similarity:	44.44%	Conservative:	7
Best Local Similarity:	31.48%	Mismatches:	17
Query Match:	27.29%	Indels:	13
DB:	11	Gaps:	1

US-10-089-278-6 COPY 1 45 (1-45) x ADM45388 (1-648)

Qy	2	AspCysLeuThrAsnLeuAraSerAlaGluGlyLysValAspGlnAlaSerIyslleLeu	21
			:::
			:::
Db	231	GATTGTTTGCACAGGCTTATTATCACGACTGTGGCTCTTAACCGAGAAAGTGAAGTGCCA	172
Qy	22	IleLeuLeuValalaTrp-----	Trp 28
		:::::	
Db	171	CTGGTACTGGTGGCTTTGAGGTGGACGCGAGGTCCTGCACCGAGAAAGCTCTGG	112
Qy	29	GlyPheGlyThrThrAlaGluValSerThrAlaArgAlaIa	42
		:::	
Db	111	TGGCGGAATAGAAGCCCTTGTGGCCATACAGAGCTGTCTGA	70

Search completed: August 25, 2005, 15:48:01
Job time : 400 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 25, 2005, 15:41:25 ; Search time 145 Seconds

(without alignments)
507.810 Million cell updates/sec

Title: US-10-089-278-6_COPY_1_45

Perfect score: 229
Sequence: 1 MDCLNLRSAEGKVDQASKI.....AWMGFTTAETVSTABAAQFA 45

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2_1/USPTO_spool_h/US10089278/runat_25082005_131344_29690/app_query.fasta_1.199
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -DOOPCL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10089278 @CNC 1.1.69 @runat_25082005_131344_29690 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DSV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	200	87.3	4776	US-09-555-352-1	Sequence 1, Appl
2	191	83.4	220	US-08-933-616-2	Sequence 2, Appl
3	191	83.4	220	US-09-135-121B-4	Sequence 4, Appl
4	183	79.9	1704	US-08-007-282B-1	Sequence 1, Appl
5	62.5	27.3	97195	US-09-949-016-12212	Sequence 12212, A
6	62.5	27.3	97196	US-09-949-016-16971	Sequence 16971, A
7	62	27.1	423	US-08-957-001B-6	Sequence 6, Appl
8	62	27.1	423	US-09-496-301-6	Sequence 6, Appl
9	61	26.6	22123	US-09-949-016-12376	Sequence 12376, A
10	61	26.6	22124	US-09-949-016-14666	Sequence 14666, A
11	59.5	26.0	1956	US-09-724-797-57	Sequence 57, Appl
12	58.5	25.5	242	US-09-270-767-8117	Sequence 8117, Ap

13	58.5	25.5	242	4	US-09-270-767-23399	Sequence 23399, A
14	58.5	25.5	417	1	US-08-398-613A-21	Sequence 21, Appl
15	58.5	25.5	417	1	US-08-398-612A-21	Sequence 21, Appl
16	58.5	25.5	417	1	US-08-398-611A-21	Sequence 21, Appl
17	58.5	25.5	417	1	US-08-396-851A-21	Sequence 21, Appl
18	58.5	25.5	417	2	US-08-491-334A-21	Sequence 21, Appl
19	58.5	25.5	417	3	US-09-027-449-18	Sequence 18, Appl
20	58.5	25.5	417	3	US-08-804-444A-18	Sequence 18, Appl
21	58.5	25.5	417	3	US-09-026-985-18	Sequence 18, Appl
22	58.5	25.5	417	3	US-09-121-952A-18	Sequence 18, Appl
23	58.5	25.5	417	3	US-09-234-340A-18	Sequence 18, Appl
24	58	25.3	26140	4	US-09-949-016-14254	Sequence 14254, A
25	58	25.3	26140	4	US-09-949-016-14255	Sequence 14255, A
26	57	24.9	125	4	US-09-830-748B-18	Sequence 18, Appl
27	57	24.9	434	4	US-08-961-309-78	Sequence 78, Appl
28	57	24.9	434	4	US-09-830-748B-43	Sequence 43, Appl
29	57	24.9	434	4	US-09-830-748B-44	Sequence 44, Appl
30	57	24.9	524	5	PCT-US91-02942-4	Sequence 405, App
31	56.5	24.7	2069	4	US-09-799-451-405	Sequence 25928, A
32	56	24.5	422	4	US-09-270-767-25928	Sequence 6955, Ap
33	56	24.5	456	4	US-09-252-991A-5913	Sequence 6913, Ap
34	56	24.5	1116	4	US-09-252-991A-6913	Sequence 6829, Ap
35	56	24.5	1251	4	US-09-252-991A-6829	Sequence 17, Appl
36	56	24.5	1797	1	US-08-442-542-17	Sequence 17, Appl
37	56	24.5	1797	3	US-08-765-469-17	Sequence 10514, A
38	56	24.5	1926	4	US-09-270-767-10514	Sequence 16487, A
39	56	24.5	12082	4	US-09-949-016-16487	Sequence 12671, A
40	56	24.5	51698	4	US-09-949-016-12671	Sequence 32, Appl
41	56	24.5	118999	4	US-08-791-105B-32	Sequence 12672, A
42	56	24.5	134140	4	US-09-949-016-12424	Sequence 12424, A
43	56	24.5	134241	4	US-09-949-016-12424	Sequence 15813, A
44	56	24.5	134242	4	US-09-949-016-15813	Sequence 15814, A
45	56	24.5	134242	4	US-09-949-016-15814	Sequence 15814, A

ALIGNMENTS

RESULT 1
US-09-555-352-1
; Sequence 1, Application US/09555352
; Patent No. 6544779
; GENERAL INFORMATION:
; APPLICANT: Cichutek, Klaus
; APPLICANT: Merget-Millitzer, Heike
; TITLE OF INVENTION: PSEUDO-TYPE RETROVIRAL VECTORS WITH
; TITLE OF INVENTION: MODIFIABLE SURFACE CAPSID PROTEINS
; FILE REFERENCE: 11692-005001
; CURRENT APPLICATION NUMBER: US/09/555,352
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: PCT/DE98/03542
; PRIOR FILING DATE: 1998-11-27
; PRIOR APPLICATION NUMBER: DE 197 52 855.4
; PRIOR FILING DATE: 1997-11-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4776
; TYPE: DNA
; ORGANISM: Murine leukemia virus
US-09-555-352-1

Alignment Scores:
Pred. No.: 4-82e-21 Length: 4776
Score: 200.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 87.34% Indels: 0
DB: 4 Gaps: 0
US-10-089-278-6_COPY_1_45 (1-45) x US-09-555-352-1 (1-4776)

QY 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyValAspClnAlaserlysville 20

REFERENCE/DOCKET NUMBER: 1644-104P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-949-016-12212

Alignment Scores:
Pred. No.: 5 8e-19 Length: 1704
Score: 185.00 Matches: 35
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.91% Indels: 0
DB: 1 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x US-08-007-282B-1 (1-1704)

Qy 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyValAspGlnAlaSerLysIle 20
Db 1 ATGGACTGCTCACCACCTCCGATCCGCTGAGGGTAAAGTTGACCCAGCGGACCAATC 60

Qy 21 LeuIleLeuValAlaTrpTrpGlyPheGlyThrAlaGlu 35

Db 61 CTAATTCTCTTGTGGCTGGTGGGGTTTGGGACCACCTGCCGAA 105

RESULT 5

US-09-949-016-12212
Sequence 12212, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12212

Alignment Scores:
Pred. No.: 1.42e+03 Length: 97195
Score: 62.50 Matches: 16
Percent Similarity: 53.85% Conservative: 5
Best Local Similarity: 41.03% Mismatches: 17
Query Match: 27.29% Indels: 1
DB: 4 Gaps: 1

US-10-089-278-6_COPY_1_45 (1-45) x US-09-949-016-12212 (1-97195)
US-09-949-016-12212
Sequence 12212, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12212

Alignment Scores:
Pred. No.: 1.42e+03 Length: 97195
Score: 62.50 Matches: 16
Percent Similarity: 53.85% Conservative: 5
Best Local Similarity: 41.03% Mismatches: 17
Query Match: 27.29% Indels: 1
DB: 4 Gaps: 1

Qy 8 ArgSerAlaGluGlyValAspGlnAlaSerLysIleLeuIleLeuValAlaTrp 27
Db 69972 CGCTCAGCAGGCGTGCGAGTCACTCAGGCTGCTGGCCTTCTGCTGCAAGTGG 70031
Qy 28 TrpGlyPheGlyThrAlaGluValSerThr---AlaArgAlaAlaGlnProAla 45
Db 70032 TGGGGGCTGAGCCAGAGCCCCCTTGGGGGCGAGGTGTGCGATGGGCTCTTCTCTGCC 70088

RESULT 6

US-09-949-016-16971
Sequence 16971, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16971

Alignment Scores:
Pred. No.: 1.42e+03 Length: 97196
Score: 62.50 Matches: 16
Percent Similarity: 53.85% Conservative: 5
Best Local Similarity: 41.03% Mismatches: 17
Query Match: 27.29% Indels: 1
DB: 4 Gaps: 1

US-10-089-278-6_COPY_1_45 (1-45) x US-09-949-016-16971 (1-97196)
US-09-949-016-16971
Sequence 16971, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16971

Qy 8 ArgSerAlaGluGlyValAspGlnAlaSerLysIleLeuIleLeuValAlaTrp 27
Db 69972 CGCTCAGCAGGCGTGCGAGTCACTCAGGCTGCTGGCCTTCTGCTGCAAGTGG 70031
Qy 28 TrpGlyPheGlyThrAlaGluValSerThr---AlaArgAlaAlaGlnProAla 45
Db 70032 TGGGGGCTGAGCCAGAGCCCCCTTGGGGGCGAGGTGTGCGATGGGCTCTTCTCTGCC 70088

RESULT 7

US-08-957-001B-6
Sequence 6, Application US/08957001B
Patent No. 6228621

GENERAL INFORMATION:
APPLICANT: Williams, William V.
APPLICANT: Madaio, Michael
APPLICANT: Weiner, David B.
TITLE OF INVENTION: IMPROVED VACCINES
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESS: Woodcock Washburn Kurtz Mackiewicz & No. 6228621ris
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103

COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: windows
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/957,001B
; FILING DATE: 23-OCT-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,592
; FILING DATE: 23-OCT-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: UPN-3303
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..427
US-08-957-001B-6

Alignment Scores:
Pred. No.: 1.12 Length: 423
Score: 62.00 Matches: 14
Percent Similarity: 54.76% Conservative: 9
Best Local Similarity: 33.33% Mismatches: 19
Query Match: 27.07% Indels: 0
DB: 3 Gaps: 0
```

US-10-089-278-6_COPY_1_45 (1-45) x US-08-957-001B-6 (1-423)

```
QY 4 LeuThrasLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIleLeuLeu 23
Db 247 CTCAGCAGCTGCATCTGACGACTCTGGGCTCTATTATTGTCAGAGAACGTCCTCG 306
QY 24 LeuValAlaTrpTrpGlyPheGlyThrAlaGluValSerThrAlaArgAlaGln 43
Db 307 GGGTTTGTATTGGGGCCAAAGGACTCTGGTCACTGTCTCTACAGCCAAACACAGCC 366
QY 44 ProAla 45
Db 367 CCATCG 372
```

```
RESULT 8
US-09-496-301-6
; Sequence 6, Application US/09496301
; Patent No. 6248565
; GENERAL INFORMATION:
; APPLICANT: Williams, William V.
; APPLICANT: Madaio, Michael
; APPLICANT: Weiner, David B.
; TITLE OF INVENTION: IMPROVED VACCINES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6248565ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM: floppy disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: windows
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/496,301
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/957,001
; FILING DATE: 23-OCT-1997
; APPLICATION NUMBER: US 60/029,592
; FILING DATE: 23-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: UPN-3303
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..427
US-09-496-301-6

Alignment Scores:
Pred. No.: 1.12 Length: 423
Score: 62.00 Matches: 14
Percent Similarity: 54.76% Conservative: 9
Best Local Similarity: 33.33% Mismatches: 19
Query Match: 27.07% Indels: 0
DB: 3 Gaps: 0
```

US-10-089-278-6_COPY_1_45 (1-45) x US-09-496-301-6 (1-423)

```
QY 4 LeuThrasLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIleLeuLeu 23
Db 247 CTCAGCAGCTGCATCTGACGACTCTGGGCTCTATTATTGTCAGAGAACGTCCTCG 306
QY 24 LeuValAlaTrpTrpGlyPheGlyThrAlaGluValSerThrAlaArgAlaGln 43
Db 307 GGGTTTGTATTGGGGCCAGGACTCTGGTCACTGTCTCTACAGCCAAACACAGCC 366
QY 44 ProAla 45
Db 367 CCATCG 372
```

```
RESULT 9
US-09-949-016-12376/c
; Sequence 12376, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12376
```

```
; LENGTH: 22123
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) .. (22123)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12376

Alignment Scores:
Pred. No.: 332 Length: 22123
Score: 61.00 Matches: 9
Percent Similarity: 82.35% Conservative: 5
Best Local Similarity: 52.94% Mismatches: 3
Query Match: 26.64% Indels: 0
DB: 4 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x US-09-949-016-12376 (1-22123)
QY 15 AppGlnAlaSerLysIleLeuLeuValAlaTrpGlyPheGly 31
Db 7273 GACCAAAATCACAGCTTATTCTACATCTCTGGCCTGGGGGCTGGGG 7223

RESULT 10
US-09-949-016-14666/c
; Sequence 14666, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14666
; LENGTH: 22124
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) .. (22124)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14666

Alignment Scores:
Pred. No.: 332 Length: 22124
Score: 61.00 Matches: 9
Percent Similarity: 82.35% Conservative: 5
Best Local Similarity: 52.94% Mismatches: 3
Query Match: 26.64% Indels: 0
DB: 4 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x US-09-949-016-14666 (1-22124)
QY 15 AppGlnAlaSerLysIleLeuLeuValAlaTrpGlyPheGly 31
Db 7273 GACCAAAATCACAGCTTATTCTACATCTCTGGCCTGGGGGCTGGGG 7223

RESULT 11
US-09-724-797-57
; Sequence 57, Application US/09724797
; Patent No. 6733998
; GENERAL INFORMATION:
; APPLICANT: Jon S. THORSON
; TITLE OF INVENTION: MICROMONOSPORA ECHINOSPORA GENES
```

```
; TITLE OF INVENTION: ENCODING FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: CALICHEAMICIN AND SELF-RESISTANCE THERETO
; FILE REFERENCE: 2653-40
; CURRENT APPLICATION NUMBER: US/09/724,797
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/111,325
; PRIOR FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Bacteria
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (1956)
US-09-724-797-57

Alignment Scores:
Pred. No.: 21.8 Length: 1956
Score: 59.50 Matches: 17
Percent Similarity: 48.78% Conservative: 3
Best Local Similarity: 41.46% Mismatches: 12
Query Match: 25.98% Indels: 9
DB: 4 Gaps: 2

US-10-089-278-6_COPY_1_45 (1-45) x US-09-724-797-57 (1-1956)
QY 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIle 20
Db 400 ATGGGCTGCTGCCCGGCGACCTCAACGCGGACGCGCGCGAC----- 444
QY 21 LeuIleLeuValAlaTrpGlyPheGlyThrAlaGluValSerThrAlaArg 40
Db 445 -----CTGCTGCTACTGTGGTGG-----GGCCGCGACCCCGGTGCTCTTCTGCGCCG 492

QY 41 Ala 41
Db 493 GCG 495

RESULT 12
US-09-270-767-8117
; Sequence 8117, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8117
; LENGTH: 242
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-8117

Alignment Scores:
Pred. No.: 1.88 Length: 242
Score: 58.50 Matches: 17
Percent Similarity: 48.00% Conservative: 7
Best Local Similarity: 34.00% Mismatches: 14
Query Match: 25.55% Indels: 12
DB: 4 Gaps: 2

US-10-089-278-6_COPY_1_45 (1-45) x US-09-270-767-8117 (1-242)
QY 1 MetAspCysLeu-----ThrAsnLeuArgSerAlaGluGlyLysValAsp 15
Db 5 CTTGACTGCTTGATTTCTTTAAACCATTTACCTTCGAGGCTCGAGGC----- 55
QY 16 GlnAlaSerLysIleLeuIleLeuValAlaTrpGlyPheGlyThrAlaArg 35
```


;; FILING DATE: 01-Mar-1995
;; APPLICATION NUMBER: 08/205864
;; FILING DATE: 03-MAR-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Love, Richard B.
;; REGISTRATION NUMBER: 34,659
;; REFERENCE/DOCKET NUMBER: P0874P1
;; TELEPHONE: 415/225-5530
;; TELEFAX: 415/952-9881
;; TELEX: 910/371-7168
;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 417 base pairs
;; TYPE: Nucleic Acid
;; STRANDEDNESS: Double
;; TOPOLOGY: Linear
US-08-398-612A-21

Alignment Scores:
Pred. No.: 3.9 Length: 417
Score: 58.50 Matches: 13
Percent Similarity: 51.11% Conservative: 10
Best Local Similarity: 28.89% Mismatches: 19
Query Match: 25.55% Indels: 3
DB: 1 Gaps: 1

US-10-089-278-6_COPY_1_45 (1-45) x US-08-398-612A-21 (1-417)

Qy 4 LeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIleLeuIleLeu 23
Db 272 ATGAGCAGTCTGAAGTCTGAGGACACAGCCATGTTTACTGTGCAAGGCCCTCATTAGT 331
Qy 24 LeuValAlaTrp-----TrpGlyPheGlyThrThrAlaGluValSerThrAlaArg 40
Db 332 TCGGCTACTTGGTTGGTTACTGGGGCCAGGGGACTCTGGTCACTGTCTCTGCAGCCAAA 391
Qy 41 AlaAlaGlnProAla 45
Db 392 ACAACAGCCCCATCT 406

Search completed: August 25, 2005, 19:15:06
Job time : 182 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 25, 2005, 15:31:58 ; Search time 491 Seconds
(without alignments)
599.672 Million cell updates/sec

Title: US-10-089-278-6_COPY_1_45

Perfect score: 229

Sequence: 1 MDCLNLRSAEGKVDQASKI.....AWMGFTTAEVSTARAQPA 45

Scoring table: BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DRV=xlh
-Qs/cgn2 1/USPTO spool h/US10089278/runat 25082005 131344 29714/app query.fasta_1.199
-DB=Published Applications NA -QWFT=fastap -SUFF1=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODES=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USRP=US10089278 @CGN 1.1 480 @runat 25082005 131344 29714
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCk=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/us07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:
22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq:
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	191	83.4	220	US-10-211-488-4	Sequence 4, Appli
2	66.5	29.0	1222	US-10-437-963-22150	Sequence 22150, A
3	63	27.5	978	US-09-974-300-2863	Sequence 2863, Ap
4	60.5	26.4	822	US-10-425-115-16327	Sequence 16327, A
5	60.5	26.4	924	US-10-027-632-120678	Sequence 120678, A
6	60.5	26.4	924	US-10-027-632-120678	Sequence 120678, A
7	60.5	26.4	930	US-10-027-632-249344	Sequence 249344, A
8	60.5	26.4	930	US-10-027-632-249344	Sequence 249344, A
9	60.5	26.4	9924	US-10-085-117-115	Sequence 115, App
10	60	26.2	433	US-10-357-930-51009	Sequence 51009, A
11	60	26.2	2034	US-10-388-838-34	Sequence 34, Appl
12	60	26.2	25310	US-10-388-838-33	Sequence 33, Appl
13	60	26.2	40392	US-09-954-456-44	Sequence 44, Appl
14	60	26.2	40392	US-09-954-456-687	Sequence 687, App
15	60	26.2	40392	US-10-843-641A-3071	Sequence 3071, Ap
16	60	26.2	40392	US-10-843-641A-3714	Sequence 3714, Ap
17	60	26.2	42104	US-10-087-192-1138	Sequence 1138, Ap
18	60	26.2	160482	US-10-087-192-100	Sequence 100, App
19	60	26.2	519599	US-10-737-082-73	Sequence 73, Appl
20	60	26.2	519599	US-10-765-790-73	Sequence 73, Appl
21	59.5	26.0	1956	US-10-152-886-20	Sequence 20, Appl
22	59.5	26.0	1956	US-11-053-052-20	Sequence 20, Appl
23	59.5	26.0	1956	US-11-053-052-20	Sequence 20, Appl
24	59	25.8	183	US-10-437-963-86046	Sequence 86046, A
25	59	25.8	803	US-10-027-632-8354	Sequence 8354, Ap
26	59	25.8	803	US-10-027-632-8354	Sequence 8354, Ap
27	59	25.8	2527	US-10-425-115-9391	Sequence 9391, Ap
28	59	25.8	166536	US-10-981-277-35	Sequence 35, Appl
29	58.5	25.5	417	US-09-726-258-18	Sequence 18, Appl
30	58	25.3	471	US-10-916-840-103	Sequence 103, App
31	58	25.3	657	US-10-128-520-309	Sequence 303, App
32	58	25.3	704	US-10-767-701-631	Sequence 631, App
33	58	25.3	2049	US-10-739-930-714	Sequence 714, App
34	58	25.3	2319	US-10-104-047-243	Sequence 243, App
35	58	25.3	2598	US-09-815-242-6054	Sequence 6054, Ap
36	58	25.3	2598	US-10-282-122A-20355	Sequence 20355, A
37	58	25.3	11850	US-10-893-671-46	Sequence 46, Appl
38	57.5	25.1	1628	US-10-437-963-8553	Sequence 8553, Ap
39	57.5	25.1	27413	US-10-087-192-1477	Sequence 1477, Ap
40	57	24.9	273	US-09-783-590-2174	Sequence 2174, Ap
41	57	24.9	434	US-10-255-478-78	Sequence 78, Appl
42	57	24.9	552	US-10-027-632-142553	Sequence 142553, A
43	57	24.9	552	US-10-027-632-142553	Sequence 142553, A
44	57	24.9	561	US-09-918-995-12719	Sequence 12719, A
45	57	24.9	633	US-10-027-632-91485	Sequence 91485, A

ALIGNMENTS

RESULT 1
US-10-211-488-4
; Sequence 4, Application US/10211488
; Publication No. US2003001740A1
; GENERAL INFORMATION:
; APPLICANT: Dornburg, Ralph C
; TITLE OF INVENTION: CELL-TYPE SPECIFIC GENE TRANSFER USING RETROVIRAL VECTORS CONTAINING
; FILE OF INVENTION: ANTIBODY-ENVELOPE FUSION PROTEINS AND WILD-TYPE ENVELOPE FUSION
; FILE REFERENCE: BXTG 5870.16
; CURRENT APPLICATION NUMBER: US/10/211,488
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 09/135,121
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 08/933,616
; PRIOR FILING DATE: 1997-08-28
; PRIOR APPLICATION NUMBER: US 08/205,980


```

Pred. No.: 9.95 Length: 822
Score: 60.50 Matches: 17
Percent Similarity: 52.17% Conservative: 7
Best Local Similarity: 36.96% Mismatches: 21
Query Match: 26.42% Indels: 1
DB: 20 Gaps: 1

US-10-089-278-6_COPY_1_45 (1-45) x US-10-425-115-66327 (1-822)

QY 1 MetAsepcysLeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIle 20
Db 231 ATGGGGTCTGCGGAGCAGCAGACCTGGAGCGGAGGTAGAGAGATGCGCGCGG 290
QY 21 LeuIleLeuValAlaTrp---TTPGlyPheGlyThrThrAlaGluValSerThrAla 39
Db 291 CTGCTGCTCATTGCGGCGTGGGGTGGAGCCCGCGCGCGCGCCGCGCGCGCC 350
QY 40 AtqAlaAlaGlnProAla 45
Db 351 CGCGCGCGTGGAGCCCGCC 368

RESULT 5
US-10-027-632-120678
; Sequence 120678, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120678
; LENGTH: 924
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-120678

Alignment Scores:
Pred. No.: 11.5 Length: 924
Score: 60.50 Matches: 15
Percent Similarity: 44.44% Conservative: 5
Best Local Similarity: 33.33% Mismatches: 18
Query Match: 26.42% Indels: 7
DB: 13 Gaps: 1

US-10-089-278-6_COPY_1_45 (1-45) x US-10-027-632-120678 (1-924)

QY 4 LeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIle----- 20
Db 575 TTGACAAAGCAAAAGCAAAATGTAACAAATGTAATATCATATCATGCTTGC 634
QY 21 -----LeuIleLeuValAlaTrpTTPGlyPheGlyThrThrAlaGluVal 36
Db 635 ATTTAGGCGAGTTTGTCTTTTGTGGGGTGGTGGGGTGGTGGTGGTGGTGG 694
QY 37 SerThrAlaArgAla 41
Db 695 AAGATGGAAGGAGT 709

RESULT 7
US-10-027-632-249344
; Sequence 249344, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2002-04-20
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20

```

```

Db 695 AAGATGGAAGGAGT 709

RESULT 6
US-10-027-632-120678
; Sequence 120678, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2002-04-20
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120678
; LENGTH: 924
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-120678

Alignment Scores:
Pred. No.: 11.5 Length: 924
Score: 60.50 Matches: 15
Percent Similarity: 44.44% Conservative: 5
Best Local Similarity: 33.33% Mismatches: 18
Query Match: 26.42% Indels: 7
DB: 17 Gaps: 1

US-10-089-278-6_COPY_1_45 (1-45) x US-10-027-632-120678 (1-924)

QY 4 LeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIle----- 20
Db 575 TTGACAAAGCAAAAGCAAAATGTAACAAATGTAATATCATATCATGCTTGC 634
QY 21 -----LeuIleLeuValAlaTrpTTPGlyPheGlyThrThrAlaGluVal 36
Db 635 ATTTAGGCGAGTTTGTCTTTTGTGGGGTGGTGGGGTGGTGGTGGTGGTGG 694
QY 37 SerThrAlaArgAla 41
Db 695 AAGATGGAAGGAGT 709

RESULT 7
US-10-027-632-249344
; Sequence 249344, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2002-04-20
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20

```

```
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 249344
; LENGTH: 930
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-249344
```

```
Alignment Scores:
Pred. No.: 11.6 Length: 930
Score: 60.50 Matches: 15
Percent Similarity: 44.44% Conservative: 5
Best Local Similarity: 33.33% Mismatches: 18
Query Match: 26.42% Indels: 7
DB: 13 Gaps: 1
```

US-10-089-278-6_COPY_1_45 (1-45) x US-10-027-632-249344 (1-930)

```
QY 4 LeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIle----- 20
||||| : : : : :
Db 580 TTGACACAAAGCAAAAGGCAAAATGCTGCAACAAATGAATATCATATATCAGCTTGCAA 639

QY 21 -----LeulleLeuValAlaTTPTrpGlyPheGlyThrAlaGluVal 36
||||| : : : : :
Db 640 ATTTAGGAGGTTTGTCTGCTTTTGGTGGGTTGTTGGTTCATCATCAGCTGGTG 699

QY 37 SerThrAlaArgAla 41
||||| : : : : :
Db 700 AAGATGGAAGAGGT 714
```

```
RESULT 8
US-10-027-632-249344
; Sequence 249344, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 249344
; LENGTH: 930
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-249344
```

```
Alignment Scores:
Pred. No.: 11.6 Length: 930
Score: 60.50 Matches: 15
Percent Similarity: 44.44% Conservative: 5
Best Local Similarity: 33.33% Mismatches: 18
Query Match: 26.42% Indels: 7
DB: 17 Gaps: 1

US-10-089-278-6_COPY_1_45 (1-45) x US-10-027-632-249344 (1-930)

QY 4 LeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIle----- 20
||||| : : : : :
Db 580 TTGACACAAAGCAAAAGGCAAAATGCTGCAACAAATGAATATCATATATCAGCTTGCAA 639

QY 21 -----LeulleLeuValAlaTTPTrpGlyPheGlyThrAlaGluVal 36
||||| : : : : :
Db 640 ATTTAGGAGGTTTGTCTGCTTTTGGTGGGTTGTTGGTTCATCATCAGCTGGTG 699

QY 37 SerThrAlaArgAla 41
||||| : : : : :
Db 700 AAGATGGAAGAGGT 714

RESULT 9
US-10-085-117-115
; Sequence 115, Application US/10085117
; Publication No. US2003023234A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115
; LENGTH: 99924
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(99924)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-115

Alignment Scores:
Pred. No.: 4.51e+03 Length: 99924
Score: 60.50 Matches: 15
Percent Similarity: 64.52% Conservative: 5
Best Local Similarity: 48.39% Mismatches: 9
Query Match: 26.42% Indels: 2
DB: 17 Gaps: 1

US-10-089-278-6_COPY_1_45 (1-45) x US-10-085-117-115 (1-99924)

QY 17 AlaSerLysIleLeulleLeuValAlaTTPTrpGlyPheGlyThrAlaGluVal 36
||||| : : : : :
Db 76699 AGTTCAAAGGTTCTTGTCTGCTGGGCTGGTGGTCTGCGGACACACCTTGAGGC 76758

QY 36 1---SerThrAlaArgAlaAlaGlnProAla 45
||||| : : : : :
Db 76759 TAAGAGTGATACAAAGGCTAGAGAGATGGCT 76789

RESULT 10
US-10-357-930-51009
; Sequence 51009, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
```

;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;; TITLE OF INVENTION: HUMAN PROSTATE CANCER
;; FILE REFERENCE: MRI-007BCN
;; CURRENT APPLICATION NUMBER: US/10/357,930
;; CURRENT FILING DATE: 2003-02-04
;; PRIOR APPLICATION NUMBER: 09/785,276
;; PRIOR FILING DATE: 2003-02-16
;; PRIOR APPLICATION NUMBER: 60/183,319
;; PRIOR FILING DATE: 2000-02-17
;; PRIOR APPLICATION NUMBER: 60/189,862
;; PRIOR FILING DATE: 2000-03-16
;; PRIOR APPLICATION NUMBER: 60/207,454
;; PRIOR FILING DATE: 2000-05-25
;; PRIOR APPLICATION NUMBER: 60/211,314
;; PRIOR FILING DATE: 2000-06-09
;; PRIOR APPLICATION NUMBER: 60/219,007
;; PRIOR FILING DATE: 2000-07-18
;; PRIOR APPLICATION NUMBER: 60/255,281
;; PRIOR FILING DATE: 2000-12-13
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 51009
;; LENGTH: 433
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 43
;; OTHER INFORMATION: n = A,T,C or G

US-10-357-930-51009

Alignment Scores:

Pred. No.: 5.31 Length: 433

Score: 60.00 Matches: 14

Percent Similarity: 64.10% Conservatives: 11

Best Local Similarity: 35.90% Mismatches: 12

Query Match: 26.20% Indels: 2

DB: 20 Gaps: 1

US-10-089-278-6_COPY_1_45 (1-45) x US-10-357-930-51009 (1-433)

QY 8 ArgSerAlaGluGlyLysValAspGlnAlaSerLysIleLeuIleLeuValAlaTrp 27

DB 276 AGAGCTTCTCAATAGAAATAGACAGCGCGAGTCATCGCTGTAATCCAGCACTT 335

QY 28 TTP-----GlyPheGlyThrThrAlaGluValSerThrAlaArgAlaGlnPro 44

DB 336 TGGAGGCGCGGCGCGCGATCAGCTGAGGTTCAGAGATTGAGACCTGCCTGGCCA 392

RESULT 11

US-10-388-838-34/c

;; Sequence 34, Application US/10388838

;; Publication No. US20040180344A1

;; GENERAL INFORMATION:

;; APPLICANT: David W. Morris

;; APPLICANT: Marc Malandro

;; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer

;; FILE REFERENCE: 529452001600

;; CURRENT APPLICATION NUMBER: US/10/388,838

;; CURRENT FILING DATE: 2003-03-14

;; NUMBER OF SEQ ID NOS: 114

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 34

;; LENGTH: 2034

;; TYPE: DNA

;; ORGANISM: Mus musculus

US-10-388-838-34

Alignment Scores:

Pred. No.: 38.1 Length: 2034

Score: 60.00 Matches: 14

Percent Similarity: 47.22% Conservatives: 3

Best Local Similarity: 38.89% Mismatches: 19
Query Match: 26.20% Indels: 0
DB: 19 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x US-10-388-838-34 (1-2034)

QY 10 AlaGluGlyLysValAspGlnAlaSerLysIleLeuIleLeuValAlaTrpTrpGly 29

DB 1220 GCATCGGGTTCCTCCCGGGGGGTTCACAGCTGGTGTCTATTGGCTAGCTGGGGGGGT 1161

QY 30 PheGlyThrThrAlaGluValSerThrAlaArgAlaGlnProAla 45

DB 1160 GGAGGGGAGCCACCGTCAGCTCAACAGCAGCAGTGGCCTCGCCAGCG 1113

RESULT 12

US-10-388-838-33/c

;; Sequence 33, Application US/10388838

;; Publication No. US20040180344A1

;; GENERAL INFORMATION:

;; APPLICANT: David W. Morris

;; APPLICANT: Marc Malandro

;; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer

;; FILE REFERENCE: 529452001600

;; CURRENT APPLICATION NUMBER: US/10/388,838

;; CURRENT FILING DATE: 2003-03-14

;; NUMBER OF SEQ ID NOS: 114

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 33

;; LENGTH: 25310

;; TYPE: DNA

;; ORGANISM: Mus musculus

;; FEATURE:

;; NAME/KEY: misc_feature

;; LOCATION: (1)-(25310)

;; OTHER INFORMATION: n = A,T,C or G

US-10-388-838-33

Alignment Scores:

Pred. No.: 947 Length: 25310

Score: 60.00 Matches: 14

Percent Similarity: 47.22% Conservatives: 3

Best Local Similarity: 38.89% Mismatches: 19

Query Match: 26.20% Indels: 0

DB: 19 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x US-10-388-838-33 (1-25310)

QY 10 AlaGluGlyLysValAspGlnAlaSerLysIleLeuIleLeuValAlaTrpTrpGly 29

DB 11220 GCATCGGGTTCCTCCCGGGGGGTTCACAGCTGGTGTCTATTGGCTAGCTGGGGGGGT 1161

QY 30 PheGlyThrThrAlaGluValSerThrAlaArgAlaGlnProAla 45

DB 11160 GGAGGGGAGCCACCGTCAGCTCAACAGCAGCAGTGGCCTCGCCAGCG 11113

RESULT 13

US-09-954-456-44

;; Sequence 44, Application US/09954456

;; Patent No. US20020115057A1

;; GENERAL INFORMATION:

;; APPLICANT: Young, Paul

;; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc

;; FILE REFERENCE: 689290-76

;; CURRENT APPLICATION NUMBER: US/09/954,456

;; CURRENT FILING DATE: 2001-09-18

;; PRIOR APPLICATION NUMBER: US/60/233,617

;; PRIOR FILING DATE: 2000-09-18

;; PRIOR APPLICATION NUMBER: US/60/234,052

;; PRIOR FILING DATE: 2000-09-20

;; PRIOR APPLICATION NUMBER: US/60/234,923

;; PRIOR FILING DATE: 2000-09-25

;; PRIOR APPLICATION NUMBER: US/60/235,134

```
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 40392
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-44

Alignment Scores:
Pred. No.: 1.72e+03 Length: 40392
Score: 60.00 Matches: 11
Percent Similarity: 64.29% Conservative: 7
Best Local Similarity: 39.28% Mismatches: 10
Query Match: 26.20% Indels: 0
DB: 9 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x US-09-954-456-44 (1-40392)

Qy 17 AlaSerLySIlleLeuIleLeuValAlaTrpTrpGlyPheGlyThrThraLaGluVal 36
|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5529 GCATCTAAGCTGTACTGACCATGACCACATACATCTGGGGGTACATAGGCACGACCTGTG 5588
|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 37 SerThrAlaArgAlaAlaGlnPro 44
|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5589 CCTACCTAGGAGCTCAGGCCA 5612
|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
US-10-843-641A-3071
; Sequence 3071, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3071
; LENGTH: 40392
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-3071

Alignment Scores:
Pred. No.: 1.72e+03 Length: 40392
Score: 60.00 Matches: 11
Percent Similarity: 64.29% Conservative: 7
Best Local Similarity: 39.28% Mismatches: 10
Query Match: 26.20% Indels: 0
DB: 21 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x US-10-843-641A-3071 (1-40392)
```

Qy 17 AlaSerLysIleLeuIleLeuLeuValAlaIleTrpTrpGlyPheGlyThrThrAlaGluVal 36
 Db 5529 GCATCTAAGCTGTACCTGACCATGACCATACACTGCGGGGTACATAGGCACAGCACCTGTG 5588
 Qy 37 SerThrAlaArgAlaAlaGlnPro 44
 Db 5589 CCCTACCTAGGAGCTCACAGCCA 5612

Search completed: August 25, 2005, 19:12:23
 Job time : 520 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 25, 2005, 15:31:38 ; Search time 8866 Seconds
(without alignments)
193.198 Million cell updates/sec

Title: US-10-089-278-6_COPY_1_45

Perfect score: 229
Sequence: 1 MDCLNLSAEKQVQASKI.....AWMGFTTAETVSTARRAQA 45

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q/cgn2_1/USPTO.spool_h/US10089278/runat_25082005_131343_29665/app_query.fasta_1.199
-DB=EST -Qfmt=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcO -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10089278 @CGN 1 1 3437 @runat_25082005_131343_29665 -ICPU=3
-NO WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_g881: *
9: gb_g882: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	67.5	29.5	615	9	CC525609 CH240_400
C 2	67	29.3	580	2	BB636721 BB636721
C 3	67	29.3	742	7	CK106136 UB10CPD12
C 4	66	28.8	450	8	AQ736187 HS_2261.A
C 5	66	28.8	527	1	AJ398462 AJ398462
C 6	65.5	28.6	722	8	BZ750899 PUDCE52TB
C 7	64	27.9	399	8	BH103868 RPCI-24-3
C 8	64	27.9	529	6	CA576389 K0644A12-
C 9	64	27.9	546	6	CA576594 K0647B06-

ALIGNMENTS

RESULT 1
CC525609/c

LOCUS
DEFINITION

CC525609
genomic survey sequence.

VERSION
CC525609.1 GI:31843897

KEYWORDS
GSS.

SOURCE
Bos taurus (cow)

ORGANISM
Bos taurus

REFERENCE
1 (bases 1 to 615)

AUTHORS
Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A., L.,

Tsal, M., Cloutier, A., Lee, D., Gilm, N., Olson, T., Mayo, M.,

Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,

Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,

Schein, J., Marra, M., de Jong, P., Keete, J.W. and Kappes, S.M.

Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478

Unpublished (2003)

Other GSSs: CH240_400A7.TARBAC13P2

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240

(http://www.chori.org/bacpac/bovine240.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

CC525609 615 bp DNA linear GSS 17-JUN-2003
CH240_400A7.T7 CHORI-240 Bos taurus genomic clone CH240_400A7,

genomic survey sequence.

VERSION
CC525609.1 GI:31843897

KEYWORDS
GSS.

SOURCE
Bos taurus (cow)

ORGANISM
Bos taurus

REFERENCE
1 (bases 1 to 615)

AUTHORS
Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A., L.,

Tsal, M., Cloutier, A., Lee, D., Gilm, N., Olson, T., Mayo, M.,

Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,

Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,

Schein, J., Marra, M., de Jong, P., Keete, J.W. and Kappes, S.M.

Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478

Unpublished (2003)

Other GSSs: CH240_400A7.TARBAC13P2

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240

(http://www.chori.org/bacpac/bovine240.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

BF181593 601808995
BF580337 602097080
CC207265 CH261-56N
BB68127 BB68127
CF268107 TGESTzy1
BF160026 601767712
BG135078 P01.0.255
BB436098 BB436098
BF814911 IL5-C1014
AW086963 ga10a04.Y
AZ449569 1M0247D20
CK720205 20592 SWO
CN212563 26056 Sub
AK086483 Mus muscu
CN704703 E0531E07-
CF902621 A0343D03-
AV637896 AV637896
CF902621 A0343D03-
CO900109 Mdb55020p
CN706692 E0520F05-
CD550272 B0309D10-
CN683219 E0175F11-
CN673148 A0930B04-
CF366262 K0348C06-
CN706198 E0513F04-
CN680801 E0142C10-
CD540582 B0224C03-
CN679505 E0125D02-
CN680492 E0138C03-
CA567128 K0410B10-
CN671694 A0909E08-
CF892477 A0110C02-
CF162965 B0719B04-
CF167039 B0779C04-
CF304485 A0422H02-
BX524128 BX524128
CN679830 E0129E09-

934 2 BF181593
672 2 BF580337
1239 8 CC207265
352 2 BB68127
524 7 CF268107
971 2 BF160026
296 4 BG135078
296 2 BB436098
335 2 BF814911
437 2 AW086963
581 8 AZ449569
661 7 CK720205
672 7 CN212563
2003 3 AK086483
321 7 CN704703
330 7 CF902621
381 1 AV637896
438 7 CO900109
458 7 CN706692
500 6 CD550272
505 7 CN683219
510 7 CN673148
512 6 CF366262
515 7 CN706198
529 7 CN680801
531 6 CD540582
549 7 CN679505
553 7 CN680492
556 6 CA567128
558 7 CN671694
565 7 CF892477
566 7 CF162965
575 7 CF167039
580 7 CF304485
585 5 BX524128
590 7 CN679830

10 64 27.9
11 63 27.5
12 63 27.5
13 62.5 27.3
14 62.5 27.3
15 62.5 27.3
16 62 27.1
17 62 27.1
18 62 27.1
19 62 27.1
20 62 27.1
21 62 27.1
22 62 27.1
23 62 27.1
24 61.5 26.9
25 61.5 26.9
26 61.5 26.9
27 61.5 26.9
28 61.5 26.9
29 61.5 26.9
30 61.5 26.9
31 61.5 26.9
32 61.5 26.9
33 61.5 26.9
34 61.5 26.9
35 61.5 26.9
36 61.5 26.9
37 61.5 26.9
38 61.5 26.9
39 61.5 26.9
40 61.5 26.9
41 61.5 26.9
42 61.5 26.9
43 61.5 26.9
44 61.5 26.9
45 61.5 26.9

(<http://www.chori.org/bacpac/ordering/information.htm>). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada.

Plate: 400 row: A column: 7

Seq primer: T7

Class: BAC ends.

FEATURES
Location/Qualifiers
1..615

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="breed: Hereford"

/db_xref="taxon:9913"

/clone="CH240_400A7"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

Hereford bull LI Domino 99375; CHORI-240 Bovine BAC

library (Male) produced by Pieter de Jong"

ORIGIN

Alignment Scores:
Pred. No.: 39.1 Length: 615
Score: 67.50 Matches: 18
Percent Similarity: 47.37% Conservative: 0
Best Local Similarity: 47.37% Mismatches: 11
Query Match: 29.48% Indels: 9
DB: 9 Gaps: 2

US-10-089-278-6_COPY_1_45 (1-45) x CC525609 (1-615)

Qy 3 CysLeuThrAsnLeuArgSerAlaGluGlyValAspGlnAlaSerLysIleLeuIle 22
Db 593 TGGTTCATGAGG---AGGTCCGCCGCCAGGTAAGTCGAC-----558
Qy 23 LeuLeuValAlaTrpTropGlyPheGlyThrAlaGluValSerThrAlaArg 40
Db 557 ---CTGGCAGACGTGGTGGGGCTGGGGGTGTGAGCAAGATGAGCAGGCCAGG 507

RESULT 2

BB636721/c 580 bp mRNA linear EST 26-OCT-2001
LOCUS BB636721 RIKEN full-length enriched, adult male aorta and vein Mus
DEFINITION musculus cDNA clone A530025H02 5', mRNA sequence.

ACCESSION BB636721 GI:16472536

VERSION EST

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 580)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)

CONTACT Yoshitake Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.jp, URL: <http://genome.gsc.riken.jp/>

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new
Genes. Genome Res. 10 (10): 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11): 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2): 281-289 (2001)

Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and

Hayashizaki,Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for

further details.

e mouse tissues.

FEATURES
source

Location/Qualifiers

1..580

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="A530025H02"

/sex="male"

/tissue type="aorta and vein"

/dev stage="adult"

/lab_host="DH10B"

/clone_lib="RIKEN full-length enriched, adult male aorta

and vein"

/note="Site 1: SalI; Site 2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in

RIKEN. Division of Experimental Animal Research in Riken

contributed to prepare mouse tissues. 1st strand cDNA was

primed with a primer [5'

GAGAGAGAAAGGATCCAGAGCTCTTTTTTTTTTTTTTN 3'], cDNA was

prepared by using trehalose thermo-activated reverse

transcriptase and subsequently enriched for full-length by

cap-trapper. cDNA went through one round of normalization

to Rot = 20.0 and subtraction to Rot = 459.0. Second

strand cDNA was prepared with the primer adapter of

sequence [5' GAGAGAGATTCGAGTTAATTAATTAATTCCTCCCCCCCC

3']. cDNA was cleaved with XhoI and BamHI. Vector: a

modified pBluescript KS(+) after bulk excision from Lambda

FLC I."

ORIGIN

Alignment Scores:

Pred. No.: 43 Length: 580
Score: 67.00 Matches: 18
Percent Similarity: 63.85% Conservative: 5
Best Local Similarity: 50.00% Mismatches: 8
Query Match: 29.26% Indels: 5
DB: 2 Gaps: 1

US-10-089-278-6_COPY_1_45 (1-45) x BB636721 (1-580)

Qy 10 AlaGluGlyLysValAspGln-AlaSerLysIleLeuLeuValAlaTrpG 29
Db 446 GCTGATGGTGGCATGATTCCTCGATCGAAGCTCTTCGCTTCCAGCTGGGAGG 387

Qy 29 yPheGlyThrAlaGluValSerThrAlaArgAlaAlaGlnPro 44

Db 386 C-----AGAGAGATTTTACAGCTGCAGCTTGTGCTCCG 353

RESULT 3

CK106136

LOCUS CK106136 742 bp mRNA linear EST 01-DEC-2003
DEFINITION UB10CPD12.5pr Populus active cambium cDNA library Populus tremula
cDNA clone UB10CPD12 5', mRNA sequence.
ACCESSION CK106136
VERSION CK106136.1 GI:38590461
KEYWORDS EST.
SOURCE Populus tremula
ORGANISM Populus tremula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 742)
AUTHORS Sterky,F., Bhalerao,R.R., Unneberg,P., Segerman,B., Nilsson,P.,
Brunner,A.M., Campaa,L., Jonsson-Lindvall,J., Tandré,K.,
Strauss,S.H., Sundberg,B., Gustafsson,P., Uhlen,M., Bhalerao,R.P.,
Nilsson,O., Sandberg,G., Karlsson,J., Lundberg,J. and Jansson,S.
A Populus EST resource for functional genomics
Unpublished (2003)
TITLE A Populus EST resource for functional genomics
JOURNAL
COMMENT Other_ESTs: UB10CPD12, UB10CPD12.3pr
Contact: Bo Segerman
Umea Plant Science Center, Department of Plant Physiology
Umea University
901 87 Umea, Sweden
Tel: +46 90 786 5279
Fax: +46 90 786 6676
Email: bo.segerman@plantphys.umu.se.
Location/Qualifiers
FEATURES
source 1..742
 /organism="Populus tremula"
 /mol_type="mRNA"
 /db_xref="taxon:113636"
 /clone="UB10CPD12"
 /tissue_type="Active cambium"
 /clone_lib="Populus active cambium cDNA library"
ORIGIN
Alignment Scores:
Pred. No.: 57.5 Length: 742
Score: 67.00 Matches: 15
Percent Similarity: 50.00% Conservative: 7
Best Local Similarity: 34.09% Mismatches: 22
Query Match: 29.26% Indels: 0
DB: 7 Gaps: 0
US-10-089-278-6_COPY_1_45 (1-45) x CK106136 (1-742)
QY 2 AspCysLeuThrAsnLeuArgSerAlaGluGlyValaspGlnAlaSerIleLeu 21
Db 516 GAATGTCATTGTTCCGATTAGAAAGCTGACACCGGAAAAGAGTTCATCCATGCAGATA 575
QY 22 IleLeuLeuValAlaIatrpTrpGlyPheGlyThrAlaGluValSerThrAlaArgAla 41
Db 576 CTGATTGTTGTCGAAACGGATTTGGAACTGATTTGGCTATATAGTACTAGATGGAG 635
QY 42 AlaGlnProAla 45
Db 636 CTCAATCCAGCT 647
RESULT 4
QAQ736187/c
LOCUS
DEFINITION HS_2261_A1_G07_MR CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=2261 Col=13 Row=M, genomic survey
sequence.
ACCESSION QAQ736187
VERSION QAQ736187.1 GI:5507739
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 450)
AUTHORS Mahabadi,G.G., Wallace,J.C., Smith,K., Swartzell,S.,Holzman,T.,

Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 2261 Row: M Column: 13
Seq primer: MJ3 Reverse
Class: BAC ends
High quality sequence stop: 450.
Location/Qualifiers
1. 450
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=2261 Col=13 Row=M"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in E-Coli DH10B"

ALIGNMENT SCORES

Pred. No.:	44	Length:	450
Score:	66.00	Matches:	15
Percent Similarity:	56.10%	Conservative:	8
Best Local Similarity:	36.59%	Mismatches:	18
Query Match:	28.82%	Indels:	0
DB:	8	Gaps:	0

US-10-089-278-6_COPY_1_45 (1-45) x AQ736187 (1-450)

QY 3 CysLeuThrAsnLeuArgSerAlaGluGlyValAspGlnAlaSerLysIleLeuIle 22
437 TCGGATCCCTTTCGTGATTCTCAATCCACCGGTCGATTTCATTCACCGGTGTGTGG 378
QY 23 LeuLeuValAlaTrpGlyPheGlyThrThrAlaGluValSerThrAlaArgAlaAla 42
437 ACCCTAGTTAGTGTGGGTATGGGGTGCACCTTCGTGTATGTCAAGTCCCTCGCGCC 318
QY 43 Gln 43
437 AGA 315

RESULT 5
AJ398462
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AJ398462
AJ398462 dkfz426 Gallus gallus cdna clone 6a21r1, mRNA sequence.
AJ398462
EST.
GI:7132444
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 527)
Abdrakhmanov, I., Lodygin, D., Geroth, P., Arakawa, H., Law, A.,
Plachy, J., Korn, B. and Buerstedde, J.M.
A large database of chicken functional ESTs as a resource for the
analysis of vertebrate gene expression
Genome Res. 10 (12), 2062-2069 (2000)
20568495

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE

PUBMED 11116100
COMMENT Contact: Buerstedde JM
 Cellular Immunology
 Heinrich-Pette-Institute
 Martinistr. 52, 20251 Hamburg, Germany
 Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.
FEATURES Location/Qualifiers
 source 1..527
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="CB"
 /db_xref="taxon:9031"
 /clone="6a21rl"
 /tissue_type="Bursa of Fabricius"
 /cell_type="bursal lymphocyte"
 /dev_stage="2-3 weeks old"
 /clone_lib="dkfz426"

ORIGIN
 Alignment Scores: 53.1 Length: 527
 Pred. No.: 66.00 Matches: 17
 Score: 50.00% Conservative: 6
 Percent Similarity: 36.96% Mismatches: 21
 Best Local Similarity: 28.82% Indels: 2
 Query Match: 1 Gaps: 1
 DB: 1

US-10-089-278-6_COPY_1_45 (1-45) x AJ398462 (1-527)
 Qy 1 MetAspCysLeuThrAsnLeuArgSerAlaGluGlyIysValAspGlnAlaSerLysIle 20
 Db 224 ATGCAGACGCTGTGCAGCTGGCTCCAGCTCTCGAAGCCGAGCAGATGGCAGCTCTG 283
 Qy 21 LeuIleuLeuVal-----AlaTrpTrpGlyPheGlyThrThrAlaGluValSerThr 38
 Db 284 CTGCCCCGGCTGACCTGTGAGTCTTGGTGGGGCTTGACGTTGCACACTGTGTCACACA 343
 Qy 39 AlaArgAlaAlaGlnPro 44
 Db 344 CACGTCTCTGCTCACCCC 361

RESULT 6 BZ750899 722 bp DNA linear GSS 10-MAR-2000
LOCUS PUDCE527B ZM 0.61.0 KB Zea mays genomic clone ZMMBTrA154I07,
DEFINITION genomic survey sequence.
ACCESSION BZ750899
VERSION BZ750899.1 GI:28903248
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1. (bases 1 to 722)
 Whiteglaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
 Bennetzen,J.
 Maize Genomics Consortium
 Unpublished (2003)
 Other_GSSs: PUDCE52TD
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 source 1..722
 /organism="Zea mays"
 /mol_type="genomic DNA"

BamH1 sites using MboI partially digested male C57BL/6J DNA."

```

ORIGIN
Alignment Scores:
Pred. No.: 73.3 Length: 399
Score: 64.00 Matches: 12
Percent Similarity: 68.00% Conservative: 5
Best Local Similarity: 48.00% Mismatches: 8
Query Match: 27.95% Indels: 0
DB: 8 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x BH103868 (1-399)

Qy 11 GluGlyLysValAspGlnAlaSerLysIleLeuLeuValaLtrpGlyPhe 30
Db 338 GAAGGGAAGTTGCCAGGATCCCGCTTTCTATTAGTACAGCCTGGTGCCTTT 279

```

```

Qy 31 GlyThrThralaGlu 35
Db 278 GGAACATCCAGGAG 264

RESULT 8
CA576389
LOCUS
DEFINITION
CA576389 529 bp mRNA linear EST 19-NOV-2002
K0644A12-5N NIA Mouse Hematopoietic Stem Cell (Lin-/C-Kit-/Sca-1+)
cDNA Library (Long) Mus musculus cDNA clone NIA:K0644A12
IMAGE:30072971 5', mRNA sequence.

ACCESSION
CA576389.1 GI:25121091
VERSION
EST.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 529)
Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Luo,A., Carter,M.G.,
Taub,D., Longo,D.L., Keller,J. and Ko,M.S.H.
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
(Lin-/C-Kit-/Sca-1+) cDNA Library (Long)
Unpublished (2001)
Other_ESTs: K0644A12-3
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: K0644, row: A, column: 12
Seq primer: M13 Reverse
High quality sequence stop: 529
POLYA=No.
Location/Qualifiers
1..529
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6NCr"
/db_xref="ntaEST:K0644A12-5N"
/db_xref="taxon:10090"
/clone="NIA:K0644A12 IMAGE:30072971"
/tissue_type="Hematopoietic Stem Cell
(Lin-/C-Kit-/Sca-1+)"
/dev_stage="Age approx.10 weeks old"
/lab_host="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
(Lin-/C-Kit-/Sca-1+) cDNA Library (Long)"
/notes="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Drs. Dennis Taub, Dan Longo (National
Institute on Aging, USA), Jonathan Keller (National Cancer

```

Institute, USA). Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen):
 5'-pGACTAGTCTAGATCGGAGCGCGCCCTTTT-3' from
 1.1 ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to lone-linker LL-Sal4, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pSPORT1 plasmid vector.
 The DH10B E. coli host was transformed with the ligation
 mixture by the standard chemical method. The average
 insert size is about 2.2 kb. The library was constructed
 by Yulan Piao (NIA)."

```

ORIGIN
Alignment Scores:
Pred. No.: 102 Length: 529
Score: 64.00 Matches: 12
Percent Similarity: 57.14% Conservative: 12
Best Local Similarity: 28.57% Mismatches: 18
Query Match: 27.95% Indels: 0
DB: 6 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x CA576389 (1-529)

Qy 4 LeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIleLeuLeu 23
Db 328 CTCAGCAGCCTGACCACTGAGGACTCTGCCATCTATTACTGTGCAAGAGCTGTGGTGG 387

Qy 24 LeuValAlaLtrpGlyPheGlyThrAlaGluValSerThrAlaArgAlaAlaGln 43
Db 388 GTCCTTGTCTACTGGGGCCAGGACTCTGGTCTCTCTGCAGCCAAACACACACCC 447

Qy 44 ProAla 45
Db 448 CCATCA 453

RESULT 9
CA576594
LOCUS
DEFINITION
CA576594 546 bp mRNA linear EST 19-NOV-2002
K0647B06-5N NIA Mouse Hematopoietic Stem Cell (Lin-/C-Kit-/Sca-1+)
cDNA Library (Long) Mus musculus cDNA clone NIA:K0647B06
IMAGE:30073265 5', mRNA sequence.

ACCESSION
CA576594
VERSION
EST.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 546)
Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Luo,A., Carter,M.G.,
Taub,D., Longo,D.L., Keller,J. and Ko,M.S.H.
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
(Lin-/C-Kit-/Sca-1+) cDNA Library (Long)
Unpublished (2001)
Other_ESTs: K0647B06-3
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: K0647, row: B, column: 06
Seq primer: M13 Reverse
High quality sequence stop: 546
POLYA=No.
Location/Qualifiers
1..546
/organism="Mus musculus"
/mol_type="mRNA"

```

```

/strains="C57BL/6NCr"
/db_xref="niaEST:K0647B06-5N"
/db_xref="taxon:10090"
/clone="NIA:K0647B06 IMAGE:30073265"
/tissue_type="Hematopoietic Stem Cell
(Lin-/C-Kit-/Sca-1)"
/dev_stage="Age approx.10 weeks old"
/lab_host="DHI10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
(Lin-/C-Kit-/Sca-1) cDNA Library (long)"
/notes="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Drs. Dennis Taub, Dan Longo (National
Institute on Aging, USA), Jonathan Keller (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an Oligo(dT) primer [Invitrogen]:
5'-pGATGTTCTAGATCGAGCGCGCCCTTTTCTTTT-3' from
1.1 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker L1-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pSPORT1 plasmid vector.
The DHI10B E. coli host was transformed with the ligation
mixture by the standard chemical method. The average
insert size is about 2.2 kb. The library was constructed
by Yulan Piao (NIA)."

ORIGIN
Alignment Scores:
Pred. No.: 106 Length: 546
Score: 64.00 Matches: 12
Percent Similarity: 57.14% Conservative: 12
Best Local Similarity: 28.57% Mismatches: 18
Query Match: 27.95% Indels: 0
DB: 6 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x CA576594 (1-546)
QY 4 LeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIleLeuLeu 23
Db 328 CTCAGCAGCTGACACTGAGGACTCTGCCATCTATTACTGTGCAAGACTCTGGTCTGG 387
QY 24 LeuValAlaTrpTrpGlyPheGlyThrThrAlaGluValSerThrAlaArgAlaGln 43
Db 388 GTCTTTGCTTACTGTGGGCGCAAGGACTCTGGTCACTGTCTCTGCAGCAAAACACACCC 447
QY 44 ProAla 45
Db 448 CCATCA 453

RESULT 10
BF181593 934 bp mRNA linear EST 31-OCT-2000
LOCUS BF181593
DEFINITION BF181593
ACCESSION BF181593
VERSION BF181593
KEYWORDS BF181593.1 GI:11059735
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 934)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

ORIGIN
Alignment Scores:
Pred. No.: 106 Length: 546
Score: 64.00 Matches: 12
Percent Similarity: 57.14% Conservative: 12
Best Local Similarity: 28.57% Mismatches: 18
Query Match: 27.95% Indels: 0
DB: 6 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x BF181593 (1-934)
QY 4 LeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIleLeuLeu 23
Db 195 ATGACCAACTTAAGAGCTGAAGACTCTGGCATTATTACTGTACCAAGGAGTTGG----- 248
QY 24 LeuValAlaTrpTrpGlyPheGlyThrThrAlaGluValSerThrAlaArgAlaGln 43
Db 249 -----GCTACTGGGGCGCAAGGCAACCTCTCACAGTCTCTCTCAGCAAAACACACCC 302
QY 44 ProAla 45
Db 303 CCATCT 308

RESULT 11
BF580337 672 bp mRNA linear EST 12-DEC-2000
LOCUS BF580337
DEFINITION BF580337
ACCESSION BF580337
VERSION BF580337.1 GI:11654049
KEYWORDS BF580337.1
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 672)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9320 row: m column: 23
High quality sequence stop: 730.
Location/Qualifiers
1..934
/organism="Mus musculus"
/mol_type="mRNA"
/strains="mix FVB/N, C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4039870"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DHI10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9320 row: m column: 23
High quality sequence stop: 730.
Location/Qualifiers
1..934
/organism="Mus musculus"
/mol_type="mRNA"
/strains="mix FVB/N, C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4039870"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DHI10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

```

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM9795 row: 0 column: 19
 High quality sequence stop: 665.
 Location/Qualifiers

FEATURES

source 1..672

/organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="FVB/N"
 /strain="taxon:10090"
 /clone="IMAGE:421732"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI_CGAP_Co24"
 /note="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
 Average insert size 1.6 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Alignment Scores:
 Pred. No.: 188 Length: 672
 Score: 63.00 Matches: 13
 Percent Similarity: 68.18% Conservatives: 2
 Best Local Similarity: 59.09% Mismatches: 7
 Query Match: 27.51% Indels: 0
 DB: 2 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x BF580337 (1-672)

QY 23 LeuLeuValAlaTrpTtpGlyPheGlyThrThrAlaGluValSerThrAlaArgAlaLa 42
 Db 401 CTCCTGTTCTTACTAGGGCCAAAGGACTCTGTCACCTGCTCTACAGAGCCTGCAAGA 460

QY 43 GlnPro 44

Db 461 GAGCCC 466

RESULT 12

CC207265/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2003)

Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Insert Length: 182000 Std Error: 0.00

Seq primer: Sp6 ATTTAGGTGACACTATAG

Class: BAC ends

High quality sequence start: 94

High quality sequence stop: 510.

Location/Qualifiers

1..1239

/organism="Gallus gallus"

/mol_type="genomic DNA"

/strain="Red Jungle Fowl"

/db_xref="taxon:9031"

/clone="CH261-56N10"

FEATURES

source

/sex="female"

/cell_line="UCD001, inbred 256"

/clone_lib="CH261"

/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;

CH261 Female Chicken library - for library and clone

ordering information: <http://www.chori.org/bacpac>

ORIGIN

Alignment Scores:
 Pred. No.: 389 Length: 1239
 Score: 63.00 Matches: 18
 Percent Similarity: 46.51% Conservatives: 2
 Best Local Similarity: 41.86% Mismatches: 23
 Query Match: 27.51% Indels: 0
 DB: 8 Gaps: 0

US-10-089-278-6_COPY_1_45 (1-45) x CC207265 (1-1239)

QY 3 CysLeuThrAsnLeuArgSerAlaGluGlyLysValAspGlnAlaSerLysIleLeuIle 22
 Db 592 TGTCTCTGTTTCAAAACTCCAAACGAGCGCGGTATCTCATAGAGAATTTTGCGA 533

QY 23 LeuLeuValAlaTrpTtpGlyPheGlyThrThrAlaGluValSerThrAlaArgAlaLa 42
 Db 532 TTGCAGAGCTCTGGCTTGTTCAGGAGCTACTGCTGAGGGATGCAATCAAGCAGCGCA 473

QY 43 GlnProAla 45

Db 472 GAGCCAGCT 464

RESULT 13

BB868127/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp

URL: <http://genome.gsc.riken.jp/>

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

Genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,

Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.

and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayaishizaki,Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
 e mouse tissues.

FEATURES

Location/Qualifiers

```
1..352
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G63002H23"
/sex="male"
/tissue_type="diencephalon"
/dev_stage="16 days neonate"
/clone_lib="RIKEN full-length enriched, 16 days neonate male diencephalon"
```

ORIGIN

Alignment Scores:
 Pred. No.: 103 Length: 352
 Score: 62.50 Matches: 19
 Percent Similarity: 65.85% Conservative: 8
 Best Local Similarity: 46.34% Mismatches: 13
 Query Match: 27.29% Indels: 1
 DB: 2 Gaps: 1

US-10-089-278-6_COPY_1_45 (1-45) x BB868127 (1-352)

```
QY 3 CysLeuThrAsnLeuArgSerAlaGluGlyValAsp---GlnAlaSerLysIleLeu 21
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 336 TGCTCTCTTAGTGTAGGAGTCGGTTGGGAAGGTGAACCAACGCAATCAAGATCCTA 277
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 22 IleLeuLeuValAlaTrpTrpGlyPheGlyThrAlaGluValSerThrAlaArgAla 41
: : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 276 CGTCGGGTTTGGCATCAATGTCACCTGCTTTGACTGCTTCGTTACTAGCAGCGGCT 217
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

QY 42 Ala 42
 |||
 Db 216 GCA 214

RESULT 14

CF268107 524 bp mRNA linear EST 13-AUG-2003
 LOCUS TgESTzj12e05.y1 TgMAS Tachyzoite cDNA Library Toxoplasma gondii
 DEFINITION cDNA clone TgESTzj12e05.y1 5', mRNA sequence.

ACCESSION

CF268107

VERSION

CF268107.1 GI:33630060

KEYWORDS

EST.

SOURCE

Toxoplasma gondii

Toxoplasma gondii

Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

Sarcocystidae; Toxoplasma.

1 (bases 1 to 524)

REFERENCE

AUTHORS

Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M., Clifton,S., Page,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Teagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Waterston,R. and Wilson,R.

Toxoplasma EST Project

Unpublished (2001)

Contact: Clifton, S.

Toxoplasma EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: toxowatson.wustl.edu

Contact David Sibley (toxowest@orcim.wustl.edu) for further information relating to organism, libraries, or clone availability.
 Seq primer: -40RP from Gibco
 High quality sequence stop: 524.

FEATURES

source

```
1..524
Location/Qualifiers
/organism="Toxoplasma gondii"
/mol_type="mRNA"
/strain="tachyzoite"
/db_xref="taxon:5811"
/clone="TgESTzj12e05.y1"
/dev_stage="Tachyzoite"
/lab_host="ElectroTen Blue cells (Stratagene)"
/clone_lib="TgMAS Tachyzoite cDNA Library"
/note="Vector: pBluescript II SK+; Site1: EcoRI; Site2: XhoI; The cDNA library was constructed by Keliang Tang, and Robert Cole at Washington University. cDNA was synthesized from Poly(A)+ mRNA using an oligo-d(T) primer containing a XhoI site. Following second strand synthesis, EcoRI adapters were ligated to the cDNA, and products were size-selected on sephacryl S500. The cDNA were directionally cloned into the EcoRI/XhoI prepared pBluescript II SK+ vector, and electroporated into ElectroTen Blue cells (Stratagene). The library may contain a small percentage of host or bacterial contaminants."
```

ORIGIN

Alignment Scores:
 Pred. No.: 165 Length: 524
 Score: 62.50 Matches: 15
 Percent Similarity: 47.37% Conservative: 3
 Best Local Similarity: 39.47% Mismatches: 11
 Query Match: 27.29% Indels: 9
 DB: 7 Gaps: 1

US-10-089-278-6_COPY_1_45 (1-45) x CF268107 (1-524)

```
QY 7 LeuArgSerAlaGluGlyValAspGlnAlaSerLysIleLeuLeuValAla 26
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 370 TTGAGATCTACATGTGCGCACGTGGAT-----ATATGT 402
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

QY 27 TrpTrpGlyPheGlyThrAlaGluValSerThrAlaArgAlaGlnPro 44
 :
 Db 403 TTCTGGGGTTTCGGCGGTACGCCACTTTTAAACAGCTAGCACATGCTGCT 456
 ||||| :

RESULT 15

BF160026/c

LOCUS BF160026

DEFINITION

mRNA sequence.

ACCESSION

BF160026

VERSION

BF160026.1 GI:11040133

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 971)

REFERENCE

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: LLAM9185 row: g column: 04

High quality sequence stop: 637.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 12:13:32 ; Search time 59 Seconds
(without alignments)
294.987 Million cell updates/sec

Title: US-10-089-278-6_COPY_1_45

Perfect score: 229

Sequence: 1 MDCLTNLRSAGKQVQASKI.....AMWGFQTAEVSTARAQA 45

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	229	100.0	302	4 AAB70844	Aab70844 SNV-env 1
2	229	100.0	309	4 AAB70841	Aab70841 SNV-env 1
3	229	100.0	315	4 AAB70843	Aab70843 SNV-env 1
4	229	100.0	329	4 AAB70840	Aab70840 SNV-env 1
5	229	100.0	330	4 AAB70842	Aab70842 SNV-env 1
6	200	87.3	232	2 AAY08856	Aay08856 Expressio
7	200	87.3	232	2 AAY08769	Aay08769 Expressio
8	191	83.4	40	2 AAR80070	Aar80070 Spleen ne
9	191	83.4	40	3 AAY70110	Aay70110 Eucaryoti
10	183	79.9	567	2 AAR71700	Aar71700 Spleen ne
11	65.5	28.6	2467	5 ABB91303	Abb91303 Herbicida
12	62	27.1	124	4 AAE03750	Aae03750 Murine PS
13	62	27.1	141	2 AAW56521	Aaw56521 Anti-DNA
14	62	27.1	466	4 AAE03755	Aae03755 Chimeric
15	61	26.6	141	3 AAY43859	Aay43859 Heavy cha
16	59.5	26.0	651	5 ABG97425	Abg97425 M. echino
17	59.5	26.0	651	6 ABU11992	Abu11992 M. echino
18	59	25.8	129	2 AAR25959	Aar25959 ICAM-1 in
19	58.5	25.5	130	2 AAR86318	Aar86318 Anti-IL-8
20	58.5	25.5	130	2 AAW23785	Aaw23785 Monoclonal
21	58.5	25.5	130	2 AAW31575	Aaw31575 Anti IL-8
22	58.5	25.5	130	2 AAW42318	Aaw42318 Anti IL-8
23	58.5	25.5	130	2 AAW33740	Aaw33740 Anti IL-8
24	58.5	25.5	130	2 AAW69306	Aaw69306 Murine an
25	58.5	25.5	130	2 AAW40121	Aaw40121 Murine mo

26	58.5	25.5	130	2 AAY29438	Aay29438 Murine 5.
27	58.5	25.5	130	3 AAY77743	Aay77743 Murine an
28	58.5	25.5	130	3 AAB30300	Aab30300 Murine an
29	58.5	25.5	130	6 ABU13777	Abu13777 Mouse ant
30	58.5	25.5	130	6 ABU59490	Abu59490 Mouse ant
31	58.5	25.5	130	7 AAE39073	Aae39073 Murine 5.
32	58.5	25.5	579	6 AAO30910	Aao30910 di-KS-ala
33	58	25.3	84	4 AAO13433	Aao13433 Human pol
34	58	25.3	221	6 ABR01537	Abr01537 Human ant
35	58	25.3	447	8 ADO31274	Adq31274 Humanised
36	58	25.3	447	8 ADO31276	Adq31276 Humanised
37	58	25.3	448	3 AAG31162	Aag31162 Arabidops
38	58	25.3	538	3 AAG31161	Aag31161 Arabidops
39	57	24.9	152	2 AAW03200	Aaw03200 Anti-idio
40	57	24.9	163	2 AAR15060	Aar15060 Murine an
41	57	24.9	163	2 AAR15200	Aar15200 R6-5-D6 a
42	56	24.5	115	4 AAE03748	Aae03748 Murine PS
43	56	24.5	118	4 AAE03753	Aae03753 Murine PS
44	56	24.5	151	7 ABO74780	Abo74780 Pseudomon
45	56	24.5	170	2 AAR77306	Aar77306 Mouse mon

ALIGNMENTS

RESULT 1
AAB70844
ID AAB70844 standard; protein; 302 AA.
XX
AC AAB70844;
XX
DT 11-SEP-2003 (revised)
DT 25-JUN-2001 (first entry)
XX
DE SNV-env leader/human 6C3-scfv fusion construct.

XX T lymphocyte; antibody; single chain variable antibody; scfv; human;
KW cell-targeting vector; antiviral; cytostatic; immunostimulant; AIDS;
KW gene therapy; immunization; diagnosis; T cell-associated disease; SCID;
KW acquired immune deficiency syndrome; severe combined immune deficiency;
KW T cell lymphoma; fusion construct.

XX Homo sapiens.
OS Spleen necrosis virus.
OS Chimeric.

XX Key Location/Qualifiers
FT Protein 1..45
FT Protein /label= SNV-env_leader
XX 46..302
XX /label= 6C3-scfv
PN DE19946142-A1.

XX 29-MAR-2001.

XX 27-SEP-1999; 99DE-01046142.

XX 27-SEP-1999; 99DE-01046142.

XX (BUND) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.

XX Cichutek K, Engelstaedter M;

XX WPI; 2001-246140/26.

XX N-PSDB; AAF61513.

XX Cell-targeting vector selective for T lymphocytes, useful in gene therapy
of e.g. acquired immune deficiency syndrome, encodes a single-chain
variable antibody fragment.

XX Claim 1; Fig 5; 18pp; German.

XX

CC This invention describes a novel cell-targeting vector (A) containing a
 CC DNA sequence (I) encoding a single-chain variable antibody fragment
 CC (scFv). The products of the invention have antiviral, cytostatic and
 CC immunostimulant activity and can be used in gene therapy, immunization
 CC and diagnosis particularly of T cell-associated diseases, specifically
 CC acquired immune deficiency syndrome (AIDS), severe combined immune
 CC deficiency (SCID) or T cell lymphoma. (A) target T cells, independently
 CC of the CD4 receptor, with high selectivity, 4-5 fold selectivity over
 CC human B cells, and 1000 fold selectivity over other human cells. A vector
 CC designated 7A5 encodes a 329 amino acid single-chain variable antibody
 CC fragment, fully defined in the specification. It was used to transform
 CC D17 (canine osteosarcoma cells susceptible to spleen necrosis virus
 CC (SNV)) cells, C8166 (human T lymphocyte) cells, and HeLa (human cervical
 CC carcinoma) cells. After 48 hours, the cells were stained with X-gal to
 CC determine transformation. The viral titer (infectious units/ml) was over
 CC 1 million for D17, 1 million for C8166 but less than 100 for HeLa,
 CC showing the high selectivity for human T cells. This sequence represents
 CC the SNV-env leader/human 6C3-scFv fusion construct used in the
 CC construction of novel cell targeting vectors described in the invention.
 CC (Updated on 11-SEP-2003 to standardise OS field)

XX Sequence 302 AA;

Query Match 100.0%; Score 229; DB 4; Length 302;
 Best Local Similarity 100.0%; Pred. No. 9e-24; Mismatches 0; Indels 0; Gaps 0;
 Matches 45; Conservative 0;

QY 1 MDCLTNLSAEGKVDQASKILILLVAVWVGFGTTAEVSTARAQA 45
 |||||
 Db 1 MDCLTNLSAEGKVDQASKILILLVAVWVGFGTTAEVSTARAQA 45

RESULT 2
 AAB70841
 ID AAB70841 standard; protein; 309 AA.

AC AAB70841;

XX 11-SEP-2003 (revised)
 DT 25-JUN-2001 (first entry)

XX SNV-env leader/human K6-scFv fusion construct.

XX T lymphocyte; antibody; single chain variable antibody; scFv; human;
 KW cell-targeting vector; antiviral; cytostatic; immunostimulant; AIDS;
 KW gene therapy; immunization; diagnosis; T cell-associated disease; SCID;
 KW acquired immune deficiency syndrome; severe combined immune deficiency;
 KW T cell lymphoma; fusion construct.

OS Homo sapiens.
 OS Spleen necrosis virus.
 OS Chimeric.

XX Key Location/Qualifiers
 FT Protein 1..45
 FT Protein /label= SNV-env_leader
 FT Protein 46..309
 FT Protein /label= K6-scFv

XX DE19946142-A1.

XX 29-MAR-2001.

XX 27-SEP-1999; 99DE-01046142.

XX 27-SEP-1999; 99DE-01046142.

XX (BUND) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.

XX Cichutek K, Engelstaedter M;

XX WPI; 2001-246140/26.

XX N-PSDB; AAP61510.

XX Cell-targeting vector selective for T lymphocytes, useful in gene therapy
 PT of e.g. acquired immune deficiency syndrome, encodes a single-chain
 PT variable antibody fragment.

PS Claim 1; Fig 2; 18pp; German.

XX This invention describes a novel cell-targeting vector (A) containing a
 CC DNA sequence (I) encoding a single-chain variable antibody fragment
 CC (scFv). The products of the invention have antiviral, cytostatic and
 CC immunostimulant activity and can be used in gene therapy, immunization
 CC and diagnosis particularly of T cell-associated diseases, specifically
 CC acquired immune deficiency syndrome (AIDS), severe combined immune
 CC deficiency (SCID) or T cell lymphoma. (A) target T cells, independently
 CC of the CD4 receptor, with high selectivity, 4-5 fold selectivity over
 CC human B cells, and 1000 fold selectivity over other human cells. A vector
 CC designated 7A5 encodes a 329 amino acid single-chain variable antibody
 CC fragment, fully defined in the specification. It was used to transform
 CC D17 (canine osteosarcoma cells susceptible to spleen necrosis virus
 CC (SNV)) cells, C8166 (human T lymphocyte) cells, and HeLa (human cervical
 CC carcinoma) cells. After 48 hours, the cells were stained with X-gal to
 CC determine transformation. The viral titer (infectious units/ml) was over
 CC 1 million for D17, 1 million for C8166 but less than 100 for HeLa,
 CC showing the high selectivity for human T cells. This sequence represents
 CC the SNV-env leader/human K6-scFv fusion construct used in the
 CC construction of novel cell targeting vectors described in the invention.
 CC (Updated on 11-SEP-2003 to standardise OS field)

XX Sequence 309 AA;

Query Match 100.0%; Score 229; DB 4; Length 309;
 Best Local Similarity 100.0%; Pred. No. 9.3e-24; Mismatches 0; Indels 0; Gaps 0;
 Matches 45; Conservative 0;

QY 1 MDCLTNLSAEGKVDQASKILILLVAVWVGFGTTAEVSTARAQA 45
 |||||
 Db 1 MDCLTNLSAEGKVDQASKILILLVAVWVGFGTTAEVSTARAQA 45

RESULT 3
 AAB70843

ID AAB70843 standard; protein; 315 AA.

AC AAB70843;

XX 11-SEP-2003 (revised)
 DT 25-JUN-2001 (first entry)

XX SNV-env leader/human 7E4-scFv fusion construct.

XX T lymphocyte; antibody; single chain variable antibody; scFv; human;
 KW cell-targeting vector; antiviral; cytostatic; immunostimulant; AIDS;
 KW gene therapy; immunization; diagnosis; T cell-associated disease; SCID;
 KW acquired immune deficiency syndrome; severe combined immune deficiency;
 KW T cell lymphoma; fusion construct.

OS Homo sapiens.
 OS Spleen necrosis virus.
 OS Chimeric.

XX Key Location/Qualifiers
 FT Protein 1..45
 FT Protein /label= SNV-env_leader
 FT Protein 46..315
 FT Protein /label= 7E4-scFv

XX DE19946142-A1.

XX 29-MAR-2001.

XX 27-SEP-1999; 99DE-01046142.

XX 27-SEP-1999; 99DE-01046142.

XX PA (BUND) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.
 XX PI Cichutek K, Engelstaedter M;
 XX XX WPI; 2001-246140/26.
 DR DR N-PSDB; AAF61512.
 XX XX
 PT PT Cell-targeting vector selective for T lymphocytes, useful in gene therapy
 PT of e.g. acquired immune deficiency syndrome, encodes a single-chain
 PT variable antibody fragment.
 XX XX
 PS Claim 1; Fig 4; 18pp; German.
 XX XX
 CC This invention describes a novel cell-targeting vector (A) containing a
 CC DNA sequence (I) encoding a single-chain variable antibody fragment
 CC (scFv). The products of the invention have antiviral, cytostatic and
 CC immunostimulant activity and can be used in gene therapy, immunization
 CC and diagnosis particularly of T cell-associated diseases, specifically
 CC acquired immune deficiency syndrome (AIDS), severe combined immune
 CC deficiency (SCID) or T cell lymphoma. (A) target T cells, independently
 CC of the CD4 receptor, with high selectivity, 4-5 fold selectivity over
 CC human B cells, and 1000 fold selectivity over other human cells. A vector
 CC designated 7A5 encodes a 329 amino acid single-chain variable antibody
 CC fragment, fully defined in the specification. It was used to transform
 CC D17 (canine osteosarcoma cells susceptible to spleen necrosis virus
 CC (SNV)) cells, C8166 (human T lymphocyte) cells, and HeLa (human cervical
 CC carcinoma) cells. After 48 hours, the cells were stained with X-gal to
 CC determine transformation. The viral titer (infectious units/ml) was over
 CC 1 million for D17, 1 million for C8166 but less than 100 for HeLa,
 CC showing the high selectivity for human T cells. This sequence represents
 CC the SNV-env leader/human 7A5-scFv fusion construct used in the
 CC construction of novel cell targeting vectors described in the invention.
 CC (Updated on 11-SEP-2003 to standardise OS field)
 XX XX
 SQ Sequence 315 AA;
 Query Match 100.0%; Score 229; DB 4; Length 315;
 Best Local Similarity 100.0%; Pred. No. 9.5e-24;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDCLTNLSAEGKVDQASKILLVAVWVGFGTTAEVSTARAQA 45
 Db 1 MDCLTNLSAEGKVDQASKILLVAVWVGFGTTAEVSTARAQA 45
 RESULT 4
 AAB70840
 ID AAB70840 standard; protein; 329 AA.
 AC AAB70840;
 XX XX
 DT 11-SEP-2003 (revised)
 DT 25-JUN-2001 (first entry)
 XX XX
 DE SNV-env leader/human 7A5-scFv fusion construct.
 XX XX
 KW T lymphocyte; antibody; single chain variable antibody; scFv; human;
 KW cell-targeting vector; antiviral; cytostatic; immunostimulant; AIDS;
 KW gene therapy; immunization; diagnosis; T cell-associated disease; SCID;
 KW acquired immune deficiency syndrome; severe combined immune deficiency;
 KW T cell lymphoma; fusion construct.
 XX XX
 OS Homo sapiens.
 OS Spleen necrosis virus.
 OS Chimeric.
 XX XX
 FH Key Location/Qualifiers
 FT Protein 1..45
 FT /label= SNV-env_leader
 FT Protein 46..329
 FT /label= 7A5-scFv
 XX XX

PN DE19946142-A1.
 XX XX
 PD 29-MAR-2001.
 XX XX
 PF 27-SEP-1999; 99DE-01046142.
 XX XX
 PR 27-SEP-1999; 99DE-01046142.
 XX XX
 PA (BUND) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.
 XX XX
 PI Cichutek K, Engelstaedter M;
 XX XX
 DR WPI; 2001-246140/26.
 DR N-PSDB; AAF61509.
 XX XX
 PT Cell-targeting vector selective for T lymphocytes, useful in gene therapy
 PT of e.g. acquired immune deficiency syndrome, encodes a single-chain
 PT variable antibody fragment.
 XX XX
 PS Claim 1; Fig 1; 18pp; German.
 XX XX
 CC This invention describes a novel cell-targeting vector (A) containing a
 CC DNA sequence (I) encoding a single-chain variable antibody fragment
 CC (scFv). The products of the invention have antiviral, cytostatic and
 CC immunostimulant activity and can be used in gene therapy, immunization
 CC and diagnosis particularly of T cell-associated diseases, specifically
 CC acquired immune deficiency syndrome (AIDS), severe combined immune
 CC deficiency (SCID) or T cell lymphoma. (A) target T cells, independently
 CC of the CD4 receptor, with high selectivity, 4-5 fold selectivity over
 CC human B cells, and 1000 fold selectivity over other human cells. A vector
 CC designated 7A5 encodes a 329 amino acid single-chain variable antibody
 CC fragment, fully defined in the specification. It was used to transform
 CC D17 (canine osteosarcoma cells susceptible to spleen necrosis virus
 CC (SNV)) cells, C8166 (human T lymphocyte) cells, and HeLa (human cervical
 CC carcinoma) cells. After 48 hours, the cells were stained with X-gal to
 CC determine transformation. The viral titer (infectious units/ml) was over
 CC 1 million for D17, 1 million for C8166 but less than 100 for HeLa,
 CC showing the high selectivity for human T cells. This sequence represents
 CC the SNV-env leader/human 7A5-scFv fusion construct used in the
 CC construction of novel cell targeting vectors described in the invention.
 CC (Updated on 11-SEP-2003 to standardise OS field)
 XX XX
 SQ Sequence 329 AA;
 Query Match 100.0%; Score 229; DB 4; Length 329;
 Best Local Similarity 100.0%; Pred. No. 1e-23;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDCLTNLSAEGKVDQASKILLVAVWVGFGTTAEVSTARAQA 45
 Db 1 MDCLTNLSAEGKVDQASKILLVAVWVGFGTTAEVSTARAQA 45
 RESULT 5
 AAB70842
 ID AAB70842 standard; protein; 330 AA.
 XX XX
 AC AAB70842;
 XX XX
 DT 11-SEP-2003 (revised)
 DT 25-JUN-2001 (first entry)
 XX XX
 DE SNV-env leader/human 7B2-scFv fusion construct.
 XX XX
 KW T lymphocyte; antibody; single chain variable antibody; scFv; human;
 KW cell-targeting vector; antiviral; cytostatic; immunostimulant; AIDS;
 KW gene therapy; immunization; diagnosis; T cell-associated disease; SCID;
 KW acquired immune deficiency syndrome; severe combined immune deficiency;
 KW T cell lymphoma; fusion construct.
 XX XX
 OS Homo sapiens.
 OS Spleen necrosis virus.
 OS Chimeric.

XX Key Location/Qualifiers
 FH Protein 1..45
 FT /label= SNV-env_leader
 FT Protein 46..330
 FT /label= 7B2-scFv
 XX
 XX DE19946142-A1.
 XX
 XX 29-MAR-2001.
 XX
 XX 27-SEP-1999; 99DE-01046142.
 XX
 XX 27-SEP-1999; 99DE-01046142.
 PR
 XX (BUND) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.
 XX
 XX Cichutek K, Engelstaedter M;
 XX
 XX WPI; 2001-246140/26.
 DR
 DR N-PSDB; AAF61511.
 XX
 XX Cell-targeting vector selective for T lymphocytes, useful in gene therapy
 PT of e.g. acquired immune deficiency syndrome, encodes a single-chain
 PT variable antibody fragment.
 XX
 XX Claim 1; Fig 3; 19pp; German.
 XX
 XX This invention describes a novel cell-targeting vector (A) containing a
 CC DNA sequence (i) encoding a single-chain variable antibody fragment
 CC (scFv). The products of the invention have antiviral, cytostatic and
 CC immunostimulant activity and can be used in gene therapy, immunization
 CC and diagnosis particularly of T cell-associated diseases, specifically
 CC acquired immune deficiency syndrome (AIDS), severe combined immune
 CC deficiency (SCID) or T cell lymphoma. (A) target T cells, independently
 CC of the CD4 receptor, with high selectivity, 4-5 fold selectivity over
 CC human B cells, and 1000 fold selectivity over other human cells. A vector
 CC designated 7A5 encodes a 329 amino acid single-chain variable antibody
 CC fragment, fully defined in the specification. It was used to transform
 CC D17 (canine osteosarcoma cells susceptible to spleen necrosis virus
 CC (SNV)) cells, C8166 (human T lymphocyte) cells, and Hela (human cervical
 CC carcinoma) cells. After 48 hours, the cells were stained with X-gal to
 CC determine transfection. The viral titer (infectious units/ml) was over
 CC 1 million for D17, 1 million for C8166 but less than 100 for Hela,
 CC showing the high selectivity for human T cells. This sequence represents
 CC the SNV-env leader/human 7B2-scFv fusion construct used in the
 CC construction of novel cell targeting vectors described in the invention.
 CC (Updated on 11-SEP-2003 to standardise OS field)
 XX
 XX SQ Sequence 330 AA;
 Query Match 100.0%; Score 229; DB 4; Length 330;
 Best Local Similarity 100.0%; Pred. No. 1e-23;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDCLTLNLSAEGKVDQASKILLVAVWVGFGTTAEVSTAAQA 45
 |||||
 DB 1 MDCLTLNLSAEGKVDQASKILLVAVWVGFGTTAEVSTAAQA 45
 |||||
 RESULT 6
 AAY08856
 ID AAY08856 standard; protein; 232 AA.
 XX
 XX AC AAY08856;
 XX
 XX 13-AUG-1999 (first entry)
 DT
 XX Expression construct pTC53 protein fragment 9.
 DE
 XX Pseudo-type retroviral vector; surface capsid protein; virus core;
 KW retroviral packaging cell; psi-negative expression construct; gag gene;
 KW pol gene; cell-specific transduction; cell targeting; gene therapy;
 KW

KW vaccination; diagnosis; cystic fibrosis; ADA-deficiency; HIV-1 infection;
 KW chronic granulomatosis.
 XX
 XX OS Spleen necrosis virus.
 OS Mus sp.
 OS Synthetic.
 XX
 XX WO9928488-A2.
 XX
 XX 10-JUN-1999.
 XX
 XX 27-NOV-1998; 98WO-DE003542.
 XX
 XX 28-NOV-1997; 97DE-01052855.
 PR
 XX (BUND) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.
 XX
 XX Cichutek K, Merget-Millitzer H;
 XX
 XX WPI; 1999-358132/30.
 DR
 DR N-PSDB; AAX77617.
 XX
 XX Pseudo-type retroviral vectors with modified surface capsid proteins.
 PT
 XX Disclosure; Fig 4A-B; 41pp; German.
 XX
 XX This invention describes novel pseudo-type retroviral vectors with
 CC modified surface capsid proteins. The vectors of the invention consist
 CC essentially of a virus core chosen from the group of murine leukemia
 CC virus (MLV), human immunodeficiency virus (HIV), simian immunodeficiency
 CC virus (SIV), lentivirus or Foamyvirus and a virus capsid protein from
 CC spleen necrotic virus (SNV). The invention also describes a retroviral
 CC packaging cell for the retroviral vector above, and also transformed with
 CC one or more psi-negative expression constructs, the gag and pol gene
 CC products of MLV, HIV, SIV or foamyvirus, or also with a psi-negative SNV-
 CC env expression construct and/or psi-negative SNV-ENV foreign polypeptide-
 CC SNV-HIV-ENV or SNV-SIV-ENV expression construct. The pseudo-type
 CC retroviral vectors with modified surface capsid proteins are suitable for
 CC cell-specific transduction of a selected mammal cell type (cell
 CC targeting). The methods are useful for the production of the pseudo-type
 CC retroviral vectors and for gene transfer in selected cell types. The
 CC vectors can be used in medicaments for gene therapy, vaccination or
 CC diagnosis. They are particularly useful for therapy of cystic fibrosis,
 CC ADA-deficiency, chronic granulomatosis or HIV-1 infection. This sequence
 CC represents protein fragments of the expression construct pTC53 which is
 CC composed from the SNV ENV protein and a murine derived scFv fragment
 XX
 XX SQ Sequence 232 AA;
 Query Match 87.3%; Score 200; DB 2; Length 232;
 Best Local Similarity 100.0%; Pred. No. 8.5e-20;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDCLTLNLSAEGKVDQASKILLVAVWVGFGTTAEVSTA 39
 |||||
 DB 1 MDCLTLNLSAEGKVDQASKILLVAVWVGFGTTAEVSTA 39
 |||||
 RESULT 7
 AAY08769
 ID AAY08769 standard; protein; 232 AA.
 XX
 XX AC AAY08769;
 XX
 XX 13-AUG-1999 (first entry)
 DT
 XX Expression construct pTC53 protein fragment 9.
 DE
 XX Cell-specific retroviral vector; antibody domain; vaccination; scFv;
 KW cell-specific transduction; B cell RNA; variable region; heavy chain;
 KW light chain; immunoglobulin; psi-negative; retroviral Env protein;
 KW capsid protein; cell targeting; gene therapy; diagnosis; cystic fibrosis;
 KW ADA-deficiency; chronic granulomatosis; HIV-1 infection.

XX OS Spleen necrosis virus.
 OS Mus sp.
 OS Synthetic.
 XX WO9928489-A2.
 XX 10-JUN-1999.
 XX 27-NOV-1998; 98WO-DE003543.
 XX 28-NOV-1997; 97DE-01052854.
 XX (BUND) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.
 XX Cichutek K, Engelstaedter M;
 XX WPI; 1999-371131/31.
 XX N-PSDB; AAX77614.
 XX Cell-specific retroviral vectors with antibody domains suitable for cell-specific transduction of selected mammal cell types - useful for vaccination and gene therapy for treatment of, e.g. cystic fibrosis.
 XX Disclosure; Fig 4A-B; 38pp; German.
 XX This invention describes the construction of novel cell-specific retroviral vectors with antibody domains suitable for cell-specific transduction of selected mammal cell types. The invention describes a method to produce cell-specific retroviral vectors which consists essentially of the following steps: (a) immunization of a mammal with one or more cell populations (b) isolation of RNA from the immunized mammal, especially the B cell RNA (c) production of a cDNA strand of the variable region of the heavy and light chains of the immunoglobulins isolated from the RNA by RT-PCR with primers for the respective immunoglobulin chains, where the primer nucleic acid sequences are for an oligopeptide linker (d) ligation of the cDNA strain to scfv-cDNA (e) ligation of the scfv cDNA in a phagemid vector and transformation of a host bacterium with the vector (f) isolation of phage, by selection of phage that bind to the cell population used in step (a) (g) cleavage of the scfv coding DNA fragments from the cell-specific phage and ligation into a psi-negative retroviral Env-expression vector (h) transformation of a Env-Scfv expression vector to be maintained in a packaging cell and (i) isolation of a packaging cell with the retroviral vectors. The pseudo-type retroviral vectors with modified surface capsid proteins are suitable for cell-specific transduction of a selected mammal cell type (cell targeting). The methods are useful for the production of the pseudo-type retroviral vectors and for gene transfer in selected cell types. The vectors can be used in medicaments for gene therapy, vaccination or diagnosis. They are particularly useful for therapy of cystic fibrosis, ADA-deficiency, chronic granulomatosis or HIV-1 infection. This sequence represents a fragment of the expression construct pTC53 which is composed from the SNV ENV protein and a murine derived scfv fragment

XX SQ Sequence 232 AA;
 Query Match 87.3%; Score 200; DB 2; Length 232;
 Best Local Similarity 100.0%; Pred. No. 8.5e-20;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDCLTNLSAEGKVDQASKILILLVAVWVGFGTTAEVSTA 39
 |||||
 Db 1 MDCLTNLSAEGKVDQASKILILLVAVWVGFGTTAEVSTA 39
 |||||
 RESULT 8
 AAR80070
 ID AAR80070 standard; protein; 40 AA.
 XX AAR80070;
 XX 09-APR-1996 (first entry)
 XX Mastadenovirus.
 OS Spleen necrosis virus.

DE Spleen necrosis virus leader sequence protein.
 XX Antibody; scFv; targeting peptide; retroviral vector; gene therapy;
 KW adenosine deaminase deficiency; ADA; cancer; hapten dinitrophenol; B6.2;
 KW HeLa; Col-1; spleen necrosis virus; SNV.
 XX Synthetic.
 XX WO9523846-A1.
 XX 08-SEP-1995.
 XX 03-MAR-1995; 95WO-US002537.
 XX 04-MAR-1994; 94US-00205980.
 XX (UYNE-) UNIV NEW JERSEY.
 XX Dornburg RC;
 XX WPI; 1995-320563/41.
 XX N-PSDB; AAT04587.
 XX Retroviral vectors for use in cell specific gene transfer - contain antibody-envelope and wild-type envelope-fusion proteins.
 XX Example; Fig 5; 35pp; English.
 XX This sequence represents the leader sequence of the spleen necrosis virus (SNV). The sequence encoding this protein was derived from the expression vector PRD114 and is included in the expression vector pTC13. In pTC13 the DNA encoding this sequence is linked to the B6.2 sequence amplified by AAT04585 and AAT04586. B6.2 acts as a targeting peptide (Tp), and will direct the retroviral vector to a cell-surface protein expressed on the surface of various human cancers (e.g. HeLa and Col-1). By using different Tp's, that recognise different cell surface antigens, the retroviral vectors containing these Tp's can be used in a cell type specific method for introducing genes into cells. These retroviral vectors can be used in the gene therapy of human genetic diseases including, adenosine deaminase (ADA) deficiency, and in clinical trials to cure cancer. A wild type envelope can be used in addition to the altered vector, and will act as a helper molecule. The helper function enhances the infection of cells by the retroviral vector

XX SQ Sequence 40 AA;
 Query Match 83.4%; Score 191; DB 2; Length 40;
 Best Local Similarity 100.0%; Pred. No. 2.2e-19;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDCLTNLSAEGKVDQASKILILLVAVWVGFGTTAEVS 37
 |||||
 Db 1 MDCLTNLSAEGKVDQASKILILLVAVWVGFGTTAEVS 37
 |||||
 RESULT 9
 AAY70110
 ID AAY70110 standard; protein; 40 AA.
 XX AAY70110;
 XX 05-JUN-2000 (first entry)
 XX Eucaryotic gene expression vector pTC13 encoding protein.
 XX Retroviral vector; cell specific gene transfer; Spleen Necrosis Virus;
 KW SNV; antibody-envelope fusion protein; retroviral envelope protein;
 KW gene therapy; antigen binding site; single chain antibody; scFv;
 KW dinitrophenol; DNP; eucaryotic gene expression vector; pTC13; B6.2 gene;
 KW tumour cell; cell-surface protein.
 XX Mastadenovirus.
 OS Spleen necrosis virus.

OS Synthetic.
 XX WO200009730-A2.
 PN
 XX
 PD 24-FEB-2000.
 XX
 PF 10-AUG-1999; 99WO-US018141.
 XX
 PR 17-AUG-1998; 98US-00135121.
 XX
 XX (UYJE-) UNIV JEFFERSON THOMAS.
 PA
 XX Dornburg RC;
 PI
 XX WPI; 2000-224358/19.
 DR N-PSDB; AA251114.
 XX
 XX Cell specific gene transfer using retroviral vectors containing antibody-
 PT envelope fusion proteins and wild type envelope proteins.
 PT
 PS Example; Fig 5; 45pp; English.
 XX
 XX The patent discloses a novel retroviral vector, particularly Spleen
 CC Necrosis Virus (SNV) vector, having target cell specificity. The vector
 CC has a targetting envelope which is a chimeric protein consisting of an
 CC antigen binding site of an antibody (e.g. anti-DNP-scFv) or another
 CC peptide that binds to a specific cell surface protein, fused to the
 CC carboxy terminal part of the retroviral envelope protein. The presence of
 CC the wild type envelope protein serves as a helper molecule to improve or
 CC supplement a functional membrane fusion domain. The antigen binding site
 CC replaces the natural viral receptor binding site. The retroviral vector
 CC is used for cell specific gene transfer, especially in gene therapy. The
 CC invention overcomes the restricted host range limitation of retroviral
 CC vectors. The present sequence is a protein encoded by a eucaryotic gene
 CC expression vector pTC13 which is derived from another vector pRD114. The
 CC vector contains a gene fragment encoding an endoplasmic reticulum (ER)
 CC recognition signal sequence, a murine leukaemia virus promoter/enhancer
 CC sequence, an adenovirus tripartite leader sequence, an SNV leader
 CC sequence and a SV40 poly A signal sequence. The vector is used to
 CC construct a targetting envelope directed to a cell-surface protein
 CC expressed on several human tumour cells. The targetting envelope contains
 CC a single chain antibody B6.2 gene fused to SNV envelope gene
 XX
 SQ Sequence 40 AA;
 Query Match 83.4%; Score 191; DB 3; Length 40;
 Best Local Similarity 100.0%; Pred. No. 2.2e-19;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDCLTNLSAEGKVDQASKILLVAVWVGFGTTAEVS 37
 Db 1 MDCLTNLSAEGKVDQASKILLVAVWVGFGTTAEVS 37
 RESULT 10
 AAR71700
 ID AAR71700 standard; protein; 567 AA.
 AC
 AC AAR71700;
 XX
 XX 29-SEP-1995 (first entry)
 DT
 XX Spleen necrosis virus env protein.
 DE
 XX Avipoxvirus; poxvirus; reticuloendotheliosis virus; REV; vaccine;
 KW poultry.
 KW
 XX Spleen necrosis virus.
 OS
 XX US5403582-A.
 PN
 XX 04-APR-1995.
 PD
 XX

PF 21-JAN-1993; 93US-00007282.
 XX
 PR 21-JAN-1993; 93US-00007282.
 XX
 PA (JAFG) NIPPON ZEON KK.
 PA (USDA) US SEC OF AGRIC.
 XX
 PI Yanagida N, Nazerian K, Calvert JG, Witter RL;
 XX WPI; 1995-146769/19.
 DR N-PSDB; AAQ86456.
 XX
 XX New Avipoxvirus expressing glycoprotein of spleen necrosis virus - used
 PT in vaccines to protect poultry against avian reticuloendotheliosis
 PT retrovirus related diseases.
 PT
 XX Claim 3; Col 19-24; 19pp; English.
 PS
 XX A claimed fowlpox virus expresses a gene (given in AAQ86456) encoding an
 CC envelope glycoprotein (AAR71700) of the spleen necrosis virus strain of
 CC REV, the gene being inserted at position 25 or 29 of the fowlpox virus
 CC genome. The recombinant virus is used as a safe, stable, cell-free
 CC vaccine
 CC
 XX Sequence 567 AA;
 SQ
 Query Match 79.9%; Score 183; DB 2; Length 567;
 Best Local Similarity 100.0%; Pred. No. 6e-17;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDCLTNLSAEGKVDQASKILLVAVWVGFGTTAE 35
 Db 1 MDCLTNLSAEGKVDQASKILLVAVWVGFGTTAE 35
 RESULT 11
 ABB91303
 ID ABB91303 standard; protein; 2467 AA.
 XX
 AC ABB91303;
 XX
 DT 31-MAY-2002 (first entry)
 DT
 XX Herbicidally active polypeptide SEQ ID NO 514.
 DE
 XX Herbicidal; plant; agriculture; herbicide.
 KW
 XX Arabidopsis thaliana.
 OS
 XX WO200210210-A2.
 PN
 XX 07-FEB-2002.
 PD
 XX
 XX 28-AUG-2001; 2001WO-EP009892.
 PF
 XX 28-AUG-2001; 2001WO-EP009892.
 PR
 XX (FARB) BAYER AG.
 PA
 XX Tietjen K, Weidler M;
 PI
 XX WPI; 2002-269010/31.
 DR
 XX Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms.
 XX
 PS Claim 5; SEQ ID NO 514; 261pp + Sequence Listing; English.
 XX
 XX The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or

CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as
 CC herbicides

XX Sequence 2467 AA;

Query Match 28.6%; Score 65.5; DB 5; Length 2467;
 Best Local Similarity 43.2%; Pred. No. 14;
 Matches 16; Conservative 3; Mismatches 15; Indels 3; Gaps 1;

QY 3 CLTNLRASGKVDQASKILLVWVGFGTTAEVSTA 39
 ID AAE03750
 DB 1799 CAAN---ARGVDSLODILRLTLTFNHHGATADVQTA 1832

RESULT 12

AAE03750
 ID AAE03750 standard; protein; 124 AA.

XX AAE03750;

DT 07-AUG-2001 (first entry)

DE Murine PSCA Ab heavy chain variable region domain from clone 5F2.4H4.1E3.

KW Murine; prostate stem cell antigen; PSCA; cytostatic; gene therapy;
 KW glycoprotein; cancer; prostate; bladder; lung; tumour; Ab; antibody;
 KW heavy chain variable domain; VH.

XX Mus musculus.

XX Key Location/Qualifiers
 FH Region 26..35
 FT /label= CDR1

FT /note= "Complementarity determining region 1"

FT Region 50..66

FT /label= CDR2

FT /note= "Complementarity determining region 2"

FT Region 99..106

FT /label= CDR3

FT /note= "Complementarity determining region 3"

XX WO200140309-A2.

XX PN

XX PD

XX 07-JUN-2001.

XX 27-OCT-2000; 2000WO-US029603.

XX 29-OCT-1999; 99US-0162558P.

XX 16-FEB-2000; 2000US-0182872P.

XX (GETH) GENENTECH INC.

XX Devaux B, Keller G, Koeppen H, Lasky LA;

XX WPI; 2001-389954/41.

XX Novel anti-prostate stem cell antigen (PSCA) antibody that internalizes

XX on binding to PSCA on mammalian cell and inhibits growth of PSCA-
 XX expressing cancer cells in vivo, useful for killing PSCA-expressing
 XX cancer cells.

XX Claim 22; Fig 12; 112pp; English.

XX The present sequence is murine prostate stem cell antigen (PSCA) antibody
 XX (Ab) heavy chain variable region domain (VH) from hybridoma clone
 XX 5F2.4H4.1E3, Acc# 2403. PSCA is a single subunit glycoprotein that is
 XX expressed on the cell surface as a glycosylphosphatidylinositol (GPI)-
 XX anchored protein. The present invention relates to anti-PSCA antibody
 XX composition and methods of killing PSCA-expressing cancer cells. PSCA is

CC useful for inhibiting and killing the growth of PSCA-expressing cancer
 CC cells such as prostate cancer, bladder cancer or lung cancer cells.
 CC Humanised antibody conjugated to a toxin or a radioactive isotope is used
 CC for killing the cancer cells. PSCA is useful for specifically targeting
 CC PSCA-expressing tumour cells in vivo and for inhibiting or killing these
 CC cells. The antibodies are also useful for treating the above mentioned
 CC cancers and for diagnosing and staging of PSCA-expressing cancer, for
 CC purification or immunoprecipitation of PSCA from cells, and for detection
 CC and quantitation of PSCA in vitro. PSCA DNA is also useful for treating
 CC cancers by gene therapy techniques

XX Sequence 124 AA;

Query Match 27.1%; Score 62; DB 4; Length 124;
 Best Local Similarity 33.3%; Pred. No. 1.4;
 Matches 14; Conservative 10; Mismatches 18; Indels 0; Gaps 0;

QY 4 LTNLRASGKVDQASKILLVWVGFGTTAEVSTARAQPA 45
 DB 83 LSSLTSEDSAVYYCALTGIVAMVWGQGTSTVTSKTTGPS 124

RESULT 13

AAW56521

ID AAW56521 standard; protein; 141 AA.

XX AC AAW56521;

XX DT 28-SEP-1998 (first entry)

XX Anti-DNA IL/IM (H221) VH-DH-JH sequence.

KW Intracellular targeting sequence; DNA vaccine; genetic immunisation;
 KW systemic lupus erythematosus; antibody; H221.

XX Mus sp.

XX Key Location/Qualifiers
 FH Region 31..35
 FT /label= CDR-I

FT Region 50..66

FT /label= CDR-II

FT Region 99..106

FT /label= CDR-III

XX WO9817323-A1.

XX PD 30-APR-1998.

XX 23-OCT-1997; 97WO-US019545.

XX 23-OCT-1996; 96US-0029592P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Williams WV, Madaio M, Weiner DB;

XX WPI; 1998-261198/23.

XX N-PSDB; AAV29840.

XX Plasmid encoding immunogenic target protein - used in, e.g. protective or
 XX therapeutic vaccines against allergy, cancer, microbial infection or auto
 XX -immune disease.

XX Example 1; Page 31-32; 84pp; English.

XX This polypeptide comprises the VH-DH-JH sequence of anti-DNA IL/IM (H221)
 CC antibody. This antibody binds to murine renal mesangial cells and aortic
 CC endothelial cells, properties associated with pathogenic systemic lupus
 CC erythematosus antibodies. H221 VL and VH DNA regions (see AAV29839-40)
 CC have been amplified by PCR (see AAV29841-54) and used to generate VH and
 CC Fv coding sequences. These were cloned into a genetic immunisation vector
 CC and administered to mice for use in DNA vaccination studies. DNA

Sequence 141 AA;

Query Match 26.8%; Score 61; DB 3; Length 141;
Best Local Similarity 28.6%; Pred. NO.2,3;
Matches 12; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

[illegible]

Search completed: August 25, 2005, 15:30:46
Job time : 65 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 15:29:49 ; Search time 27 Seconds
(without alignments)
124.415 Million cell updates/sec

Title: US-10-089-278-6 COPY 1 45

Perfect score: 229
Sequence: 1 MDCLTNLRSAGKVDQASKI.....AWWGFGTTAEVSTARAQPA 45

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pcp.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pcp.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pcp.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pcp.*
- 5: /cgn2_6/ptodata/1/iaa/PCTRUS_COMB.pcp.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	200	87.3	232	4	US-09-555-352-10	Sequence 10, Appl	
2	191	83.4	40	2	US-08-933-616-3	Sequence 3, Appli	
3	191	83.4	40	4	US-09-135-121B-5	Sequence 5, Appli	
4	183	79.9	567	1	US-08-007-282B-2	Sequence 2, Appli	
5	62	27.1	124	4	US-09-698-705-6	Sequence 6, Appli	
6	62	27.1	141	3	US-08-957-001B-7	Sequence 7, Appli	
7	62	27.1	141	3	US-09-496-301-7	Sequence 7, Appli	
8	62	27.1	466	4	US-09-698-705-11	Sequence 11, Appl	
9	59.5	26.0	651	4	US-09-724-797-58	Sequence 58, Appl	
10	58.5	25.5	130	1	US-08-398-613A-22	Sequence 22, Appl	
11	58.5	25.5	130	1	US-08-398-612A-22	Sequence 22, Appl	
12	58.5	25.5	130	1	US-08-398-611A-22	Sequence 22, Appl	
13	58.5	25.5	130	2	US-08-491-334A-22	Sequence 22, Appl	
14	58.5	25.5	130	3	US-09-027-449-19	Sequence 19, Appl	
15	58.5	25.5	130	3	US-08-804-444A-19	Sequence 19, Appl	
16	58.5	25.5	130	3	US-09-026-985-19	Sequence 19, Appl	
17	58.5	25.5	130	4	US-09-121-952A-19	Sequence 19, Appl	
18	58.5	25.5	130	4	US-09-434-340A-19	Sequence 19, Appl	
19	57	24.9	163	5	PCT-US91-02942-5	Sequence 5, Appli	
20	56	24.5	115	4	US-09-698-705-4	Sequence 4, Appli	
21	56	24.5	118	4	US-09-698-705-9	Sequence 9, Appli	
22	56	24.5	151	4	US-09-252-991A-23526	Sequence 23526, A	
23	56	24.5	170	2	US-08-652-558-40	Sequence 40, Appl	
24	56	24.5	599	1	US-08-442-542-18	Sequence 18, Appl	
25	56	24.5	599	3	US-08-765-469-18	Sequence 18, Appl	
26	55	24.0	122	3	US-08-957-001B-26	Sequence 26, Appl	
27	55	24.0	122	3	US-09-496-301-26	Sequence 26, Appl	

; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 25
; SEQ ID NO 6
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-698-705-6

Query Match 27.1%; Score 62; DB 4; Length 124;
Best Local Similarity 33.3%; Pred. No. 0.092;
Matches 14; Conservative 10; Mismatches 18; Indels 0; Gaps 0;
Qy 4 LTNLRSAEGKVDQASKILLVAVWGFGTTAEVSTARAQA 45
Db 83 LSSLTSDSAVYVCALTGIIYAMAYWGQGLTVTVSSAKTTGPS 124

RESULT 6

US-08-957-001B-7
; Sequence 7, Application US/08957001B
; Patent No. 6228621
; GENERAL INFORMATION:
; APPLICANT: Williams, William V.
; APPLICANT: Madaio, Michael
; APPLICANT: Weiner, David B.
; TITLE OF INVENTION: IMPROVED VACCINES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6228621ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: windows
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/957,001B
; FILING DATE: 23-OCT-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,592
; FILING DATE: 23-OCT-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: UPN-3303
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-957-001B-7

Query Match 27.1%; Score 62; DB 3; Length 141;
Best Local Similarity 33.3%; Pred. No. 0.11;
Matches 14; Conservative 9; Mismatches 19; Indels 0; Gaps 0;
Qy 4 LTNLRSAEGKVDQASKILLVAVWGFGTTAEVSTARAQA 45
Db 83 LSSLTSDSAVYVCARERLLGFVWGQGLTVTVSTAKTTAPS 124

RESULT 7

US-09-496-301-7

; Sequence 7, Application US/09496301
; Patent No. 6248565
; GENERAL INFORMATION:
; APPLICANT: Williams, William V.
; APPLICANT: Madaio, Michael
; APPLICANT: Weiner, David B.
; TITLE OF INVENTION: IMPROVED VACCINES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6248565ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: windows
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/496,301
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/957,001
; FILING DATE: 23-OCT-1997
; APPLICATION NUMBER: US 60/029,592
; FILING DATE: 23-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: UPN-3303
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-09-496-301-7

Query Match 27.1%; Score 62; DB 3; Length 141;
Best Local Similarity 33.3%; Pred. No. 0.11; Indels 0; Gaps 0;
Matches 14; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

Qy 4 LTNLRSAEGKVDQASKILLVAVWGFGTTAEVSTARAQA 45

Db 83 LSSLTSDSAVYVCARERLLGFVWGQGLTVTVSTAKTTAPS 124

RESULT 8

US-09-698-705-11
; Sequence 11, Application US/09698705
; Patent No. 6824780
; GENERAL INFORMATION:
; APPLICANT: Devaux, B.
; APPLICANT: Keller, G.
; APPLICANT: Koeppe, H.
; APPLICANT: Lasky, L.
; TITLE OF INVENTION: Anti-Tumor Antibody Compositions and Methods of Use
; FILE REFERENCE: PI777RI
; CURRENT APPLICATION NUMBER: US/09/698,705
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/162,558
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/182,872
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 25
; SEQ ID NO 11

```
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: sequence is chimeric mouse/human
US-09-698-705-11

Query Match          27.1%; Score 62; DB 4; Length 466;
Best Local Similarity 33.3%; Pred. No. 0.51;
Matches 14; Conservative 10; Mismatches 18; Indels 0; Gaps 0;

QY 4 LTNLRSAEGKVDQASKILLVAVWGFGTTAEVSTARAQPA 45
   ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 102 LSLRLSEDSAVYCALTIYAYWYGQGISVTVSSAKTTGPS 143

RESULT 9
US-09-724-797-58
; Sequence 58, Application US/09724797
; Patent No. 6733998
; GENERAL INFORMATION:
; APPLICANT: Jon S. THORSON
; TITLE OF INVENTION: MICROMONOSPORA ECHINOSPORA GENES
; TITLE OF INVENTION: ENCODING FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: CALICHEAMICIN AND SELF-RESISTANCE THERETO
; FILE REFERENCE: 2653-40
; CURRENT APPLICATION NUMBER: US/09/724,797
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/111,325
; PRIOR FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Bacteria
US-09-724-797-58

Query Match          26.0%; Score 59.5; DB 4; Length 651;
Best Local Similarity 41.5%; Pred. No. 1.9;
Matches 17; Conservative 3; Mismatches 12; Indels 9; Gaps 2;

QY 1 MDCLTNLRSAEGKVDQASKILLVAVWGFGTTAEVSTARA 41
   ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 134 MGCLPGLNADGRD-----LLVYWW--GRTPVFLARA 165

RESULT 10
US-08-398-613A-22
; Sequence 22, Application US/08398613A
; Patent No. 5677426
; GENERAL INFORMATION:
; APPLICANT: Fong, Sherman
; APPLICANT: Hebert, Caroline Alice
; APPLICANT: Kim, Kyung Jin
; APPLICANT: Leong, Steven R.
; TITLE OF INVENTION: Anti-IL-8 Monoclonal Antibodies for Treatment of Inflammatory
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/398,613A
; FILING DATE: 01-MAR-1994
```

```
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/205864
; FILING DATE: 03-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 874P1-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-9881
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 130 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-398-613A-22

Query Match          25.5%; Score 58.5; DB 1; Length 130;
Best Local Similarity 28.9%; Pred. No. 0.34;
Matches 13; Conservative 10; Mismatches 19; Indels 3; Gaps 1;

QY 4 LTNLRSAEGKVDQASKILLVAVW---WGFGTTAEVSTARAQPA 45
   ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 83 MSSLSKSDTAMFYCARALISSATWFGYWGQGLTVTSAAKTAPS 127

RESULT 11
US-08-398-612A-22
; Sequence 22, Application US/08398612A
; Patent No. 5686070
; GENERAL INFORMATION:
; APPLICANT: Doershuk, Claire M.
; APPLICANT: Fong, Sherman
; APPLICANT: Hebert, Caroline Alice
; APPLICANT: Kim, Kyung Jin
; APPLICANT: Leong, Steven R.
; TITLE OF INVENTION: Anti-IL-8 Monoclonal Antibodies for
; TITLE OF INVENTION: Treatment of Inflammatory Disorders
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/398,612A
; FILING DATE: 01-MAR-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/398611
; FILING DATE: 01-Mar-1995
; APPLICATION NUMBER: 08/205864
; FILING DATE: 03-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0874P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 130 amino acids
```


[illegible]

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/027,449
; FILING DATE: 20-Feb-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/074,330
; FILING DATE: 22-Jan-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/038,664
; FILING DATE: 21-Feb-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P1085R3-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 130 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-09-027-449-19

Query Match 25.5%; Score 58.5; DB 3; Length 130;
Best Local Similarity 28.9%; Pred. No. 0.34;
Matches 13; Conservative 10; Mismatches 3; Indels 3; Gaps 1;

QY 4 LTNLRSAGKGVDOASKILILVAV---WGFGTTAEVSTARRAQA 45
   :::|: : : ||| ||| ||| :|:
Db 83 MSSLKSEDTAMFYCARALISSATWFGYWGQGLTVTVSAAKTTAPS 127

RESULT 15
US-08-804-444A-19
; Sequence 19, Application US/0880444A
; Patent No. 6117980
; GENERAL INFORMATION:
; APPLICANT: Gonzalez, Tania N
; APPLICANT: Leong, Steven R.
; APPLICANT: Presta, Leonard G.
; TITLE OF INVENTION: Humanized Anti-IL-8 Monoclonal Antibodies
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,444A
; FILING DATE: 21-Feb-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P1085
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 130 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-08-804-444A-19

Query Match 25.5%; Score 58.5; DB 3; Length 130;
Best Local Similarity 28.9%; Pred. No. 0.34;
Matches 13; Conservative 10; Mismatches 3; Indels 3; Gaps 1;

QY 4 LTNLRSAGKGVDOASKILILVAV---WGFGTTAEVSTARRAQA 45
   :::|: : : ||| ||| ||| :|:
Db 83 MSSLKSEDTAMFYCARALISSATWFGYWGQGLTVTVSAAKTTAPS 127

Search completed: August 25, 2005, 15:31:51
Job time : 28 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 15:30:58 ; Search time 557 Seconds
(without alignments)
31.744 Million cell updates/sec

Title: US-10-089-278-6_COPY_1_45

Perfect score: 229

Sequence: 1 MDCLTLNLSAEKGVDQASKI.....AWGFGTTAEVSTARAQPA 45

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues

Total number of hits satisfying chosen parameters: 1767149

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCR_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCRUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	191	83.4	40	14	US-10-211-488-5
2	65.5	28.6	2481	17	US-10-732-923-13967
3	65.5	28.6	2513	17	US-10-732-923-13968
4	62	27.1	124	17	US-10-937-046-6
5	62	27.1	466	17	US-10-937-046-11
6	60.5	26.4	197	16	US-10-425-115-250990
7	60	26.2	222	16	US-10-379-392-162
8	59.5	26.0	651	14	US-10-152-886-19
9	59.5	26.0	651	20	US-11-053-576-19
10	59.5	26.0	651	20	US-11-053-052-19
11	59	25.8	717	16	US-10-425-115-194054

12	59	25.8	753	17	US-10-732-923-13984	Sequence 13984, A
13	58.5	25.5	130	10	US-09-726-258-19	Sequence 19, Appl
14	58.5	25.5	579	14	US-10-310-719-32	Sequence 32, Appl
15	58	25.3	221	16	US-10-916-840-110	Sequence 110, App
16	58	25.3	221	16	US-10-128-520-174	Sequence 174, App
17	58	25.3	527	16	US-10-739-930-6258	Sequence 6258, Ap
18	56.5	24.7	316	15	US-10-425-114-38915	Sequence 38915, A
19	56.5	24.7	327	16	US-10-425-115-289801	Sequence 289801,
20	56	24.5	115	17	US-10-937-046-4	Sequence 4, Appli
21	56	24.5	118	17	US-10-937-046-9	Sequence 9, Appli
22	56	24.5	222	16	US-10-379-392-160	Sequence 160, App
23	56	24.5	222	16	US-10-379-392-163	Sequence 163, App
24	56	24.5	504	15	US-10-369-493-10465	Sequence 10465, A
25	55.5	24.2	224	16	US-10-128-520-182	Sequence 182, App
26	55	24.0	130	17	US-10-683-451-19	Sequence 19, Appl
27	54	23.6	249	15	US-10-307-724-124	Sequence 124, App
28	54	23.6	249	16	US-10-737-290-124	Sequence 124, App
29	54	23.6	282	16	US-10-737-290-143	Sequence 143, App
30	54	23.6	347	18	US-10-880-028-18	Sequence 18, Appl
31	54	23.6	347	18	US-10-880-320-18	Sequence 18, Appl
32	54	23.6	360	18	US-10-880-028-22	Sequence 22, Appl
33	54	23.6	360	18	US-10-880-320-22	Sequence 22, Appl
34	54	23.6	362	18	US-10-880-028-23	Sequence 23, Appl
35	54	23.6	362	18	US-10-880-028-26	Sequence 26, Appl
36	54	23.6	362	18	US-10-880-320-23	Sequence 23, Appl
37	54	23.6	673	16	US-10-739-930-5553	Sequence 5553, Ap
38	54	23.6	673	16	US-10-739-930-5553	Sequence 5553, Ap
39	54	23.6	772	16	US-10-425-115-225624	Sequence 225624,
40	54	23.6	869	16	US-10-437-963-156107	Sequence 156107,
41	54	23.6	872	17	US-10-732-923-13979	Sequence 13979, A
42	54	23.6	2659	14	US-10-311-879-28	Sequence 28, Appl
43	54	23.6	10498	17	US-10-470-048B-440	Sequence 440, App
44	53.5	23.4	251	10	US-09-726-258-27	Sequence 27, Appl
45	53.5	23.4	801	16	US-10-437-963-141818	Sequence 141818,

ALIGNMENTS

RESULT 1

US-10-211-488-5
; Sequence 5, Application US/10211488
; Publication No. US20030017140A1
; GENERAL INFORMATION:
; APPLICANT: Dornburg, Ralph C
; TITLE OF INVENTION: CELL TYPE SPECIFIC GENE TRANSFER USING RETROVIRAL VECTORS CONTAINING AN ANTIBODY-ENVELOPE FUSION PROTEIN AND A MILD-TYPE ENVELOPE FUSION PROTEIN
; TITLE OF INVENTION: ANTIBODY-ENVELOPE FUSION PROTEIN AND A MILD-TYPE ENVELOPE FUSION PROTEIN
; FILE REFERENCE: BXTG 5870.16
; CURRENT APPLICATION NUMBER: US/10/211,488
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 09/135,121
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 08/933,616
; PRIOR FILING DATE: 1997-08-28
; PRIOR APPLICATION NUMBER: US 08/205,980
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 07/979,619
; PRIOR FILING DATE: 1992-11-20
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ER recognition signal sequence
US-10-211-488-5

Query Match 83.4%; Score 191; DB 14; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.3e-18;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCLTLNLSAEKGVDQASKILILLVAMWFGTTAEVS 37

```


Dbb      1 MDCLTLNSAEGKVDQASKILLVAVWVGFGTTAEVS 37



RESULT 2  
US-10-732-923-13967  
; Sequence 13967, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 13967  
; LENGTH: 2481  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-732-923--13967



Query Match          28.6%; Score 65.5; DB 17; Length 2481;  
Best Local Similarity 43.2%; Pred. No. 19;  
Matches 16; Conservative 3; Mismatches 15; Indels 3; Gaps 1;



Oy       3 CLTNLSAEGKVDAQSKILLLVAVWGFGTTAEVSTA 39  
| | | : | : | : | : | : | : | : |  
Db      1794 CAAN---AKGVDDSLDQLRLTFLFNFHGATADVQT 1827



RESULT 3  
US-10-732-923-13968  
; Sequence 13968, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 13968  
; LENGTH: 2513  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-732-923--13968



Query Match          28.6%; Score 65.5; DB 17; Length 2513;  
Best Local Similarity 43.2%; Pred. No. 19;  
Matches 16; Conservative 3; Mismatches 15; Indels 3; Gaps 1;



Oy       3 CLTNLSAEGKVDAQSKILLLVAVWGFGTTAEVSTA 39  
| | | : | : | : | : | : | : | : |  
Db      1808 CAAN---AKGVDDSLDQLRLTFLFNFHGATADVQT 1841



RESULT 4  
US-10-937-046-6  
; Sequence 6, Application US/10937046  
; Publication No. US20050036942A1  
; GENERAL INFORMATION:  
; APPLICANT: Devaux, B.  
; APPLICANT: Keller, G.  
; APPLICANT: Koepfen, H.  
; TITLE OF INVENTION: Anti-Tumor Antibody Compositions and Methods of Use  
; FILE REFERENCE: PI77R1D1  
; CURRENT APPLICATION NUMBER: US/10/937,046  
; CURRENT FILING DATE: 2004-09-08


```

```
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 250990
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(197)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_160490C.1.pap
US-10-425-115-250990
```

```
Query Match      26.4%; Score 60.5; DB 16; Length 197;
Best Local Similarity 37.0%; Pred. No. 5.8;
Matches 17; Conservative 7; Mismatches 21; Indels 1; Gaps 1;
```

```
QY 1 MDCLTNLSAEGKVDQASKILLVAV-WGFGTTAEVSTARAQPA 45
DB 1 MGVLRSTQSLAEAEVEMRGALLLHGAWMRPGGGAARAVEPA 46
```

RESULT 7

```
US-10-379-392-162
; Sequence 162, Application US/10379392
; Publication No. US20040110226A1
; GENERAL INFORMATION:
; APPLICANT: Lazar, Gregory Alan
; APPLICANT: Desjarlais, John Rudolf
; APPLICANT: Marshall, Shannon Alicia
; APPLICANT: Dahiyat, Babsil I.
; TITLE OF INVENTION: ANTIBODY OPTIMIZATION
; FILE REFERENCE: A-71386-3 463077-236
; CURRENT APPLICATION NUMBER: US/10/379,392
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,843
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/384,197
; PRIOR FILING DATE: 2002-05-29
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 162
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-379-392-162
```

```
Query Match      26.2%; Score 60; DB 16; Length 222;
Best Local Similarity 47.8%; Pred. No. 7.8;
Matches 11; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 22 ILLVAVWGFGTTAEVSTARAQPA 44
DB 105 VLQFAYWGGTTTVTSAAKTPP 127
```

RESULT 8

```
US-10-152-886-19
; Sequence 19, Application US/10152886
; Publication No. US20030064491A1
; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: Farnet, Chris
; APPLICANT: Staffa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF ENEDIYNE RING
; FILE REFERENCE: 3011-3US
; CURRENT APPLICATION NUMBER: US/10/152,886
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 102
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Micromonospora echinospora calichensis
US-10-152-886-19
```

```
Query Match      26.0%; Score 59.5; DB 14; Length 651;
Best Local Similarity 41.5%; Pred. No. 29;
Matches 17; Conservative 3; Mismatches 12; Indels 9; Gaps 2;
```

```
QY 1 MDCLTNLSAEGKVDQASKILLVAVWGFGTTAEVSTARA 41
DB 134 MCCLPGDLNADGRD-----LLVYWN--GRTPVVFARA 165
```

RESULT 9

```
US-11-053-576-19
; Sequence 19, Application US/11053576
; Publication No. US20050142601A1
; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: Farnet, Chris
; APPLICANT: Staffa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF ENEDIYNE RING
; FILE REFERENCE: 3011-3US
; CURRENT APPLICATION NUMBER: US/11/053,576
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: US/10/152,886
; PRIOR FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Micromonospora echinospora calichensis
US-11-053-576-19
```

```
Query Match      26.0%; Score 59.5; DB 20; Length 651;
Best Local Similarity 41.5%; Pred. No. 29;
Matches 17; Conservative 3; Mismatches 12; Indels 9; Gaps 2;
```

```
QY 1 MDCLTNLSAEGKVDQASKILLVAVWGFGTTAEVSTARA 41
DB 134 MCCLPGDLNADGRD-----LLVYWN--GRTPVVFARA 165
```

RESULT 10

```
US-11-053-052-19
; Sequence 19, Application US/11053052
; Publication No. US20050170411A1
; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: Farnet, Chris
; APPLICANT: Staffa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF ENEDIYNE RING
; FILE REFERENCE: 3011-3US
; CURRENT APPLICATION NUMBER: US/11/053,052
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: US/10/152,886
; PRIOR FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Micromonospora echinospora calichensis
US-11-053-052-19
```

```
Query Match      26.0%; Score 59.5; DB 20; Length 651;
Best Local Similarity 41.5%; Pred. No. 29;
Matches 17; Conservative 3; Mismatches 12; Indels 9; Gaps 2;

QY 1 MDCLTNLRSAEGKVDQASKILLVAVWFGTTAEVSTARA 41
Db 134 MGCLPGDLNADGRTD-----LLVYVW--GRTPVVFLARA 165

RESULT 11
US-10-425-115-194054
; Sequence 194054, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, Thomas J.
; APPLICANT: La Rosa, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT FILING DATE: 2003-04-28
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 194054
; LENGTH: 717
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_108558C.1.pcp
US-10-425-115-194054

Query Match      25.8%; Score 59; DB 16; Length 717;
Best Local Similarity 41.9%; Pred. No. 38;
Matches 13; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

QY 9 SAEGKVDQASKILLVAVWFGTTAEVSTA 39
Db 38 TTKGVDDSLQDILRLTLTFNHNHGDTSVQTA 68

RESULT 12
US-10-732-923-13984
; Sequence 13984, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 13984
; LENGTH: 753
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(753)
; OTHER INFORMATION: unsure at all xaa locations
US-10-732-923-13984

Query Match      25.8%; Score 59; DB 17; Length 753;
Best Local Similarity 41.9%; Pred. No. 40;
Matches 13; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

QY 9 SAEGKVDQASKILLVAVWFGTTAEVSTA 39
Db 76 TTKGVDDSLQDILRLTLTFNHNHGDTSVQTA 106
```

```
RESULT 13
US-09-726-258-19
; Sequence 19, Application US/09726258
; Publication No. US20030021790A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc., Hsei, Vanessa
; APPLICANT: Koumenis, Iphigenia
; APPLICANT: Leong, Steven R.
; APPLICANT: Presta, Leonard G.
; APPLICANT: Shahrokhi, Zahra
; APPLICANT: Zapata, Gerardo A.
; TITLE OF INVENTION: ANTIBODY FRAGMENT-POLYMER CONJUGATES AND
; TITLE OF INVENTION: HUMANIZED ANTI-IL-8 MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/726,258
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/234,182
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/094003
; FILING DATE: 24-JUL-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P1085R4-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 130 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-726-258-19

Query Match      25.5%; Score 58.5; DB 10; Length 130;
Best Local Similarity 28.9%; Pred. No. 6.9;
Matches 13; Conservative 10; Mismatches 19; Indels 3; Gaps 1;

QY 4 LTNLRSAEGKVDQASKILLVAVW--WGFGTTAEVSTARAQPA 45
Db 83 MSSLKSEDTAMPYCARALISSATWFGYWGQGLVTVSAKTTAPS 127

RESULT 14
US-10-310-719-32
; Sequence 32, Application US/10310719
; Publication No. US20030166163A1
; GENERAL INFORMATION:
; APPLICANT: Gillies, Stephen
; TITLE OF INVENTION: Immunocytokines With Modulated Selectivity
; FILE REFERENCE: LEX-020
; CURRENT APPLICATION NUMBER: US/10/310,719
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,113
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/371,966
; PRIOR FILING DATE: 2002-04-12
```

Search completed: August 25, 2005, 15:41:14
Job time : 557 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 14:14:52 ; Search time 25 Seconds
(without alignments)
173.190 Million cell updates/sec

Title: US-10-089-278-6_COPY_1_45
Perfect score: 229
Sequence: 1 MDCLTNLSAEGKVDQASKI.....AWMGFTTAENVSTARAQPA 45
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	183	79.9	567	1 VCFVAS	env polyprotein -
2	161	70.3	582	1 VCVDFAR	env polyprotein -
3	67	29.3	213	2 S68213	Ig heavy chain (Ma
4	65.5	28.6	2513	2 G98536	hypothetical prote
5	58	25.3	538	2 B84759	hypothetical prote
6	56.5	24.7	122	2 S06825	Ig heavy chain V r
7	55	24.0	331	2 AE0189	ribose transport s
8	55	24.0	469	2 S37483	Ig gamma-2a chain
9	54	23.6	214	2 G83692	hypothetical prote
10	54	23.6	510	2 A12012	hypothetical prote
11	54	23.6	904	2 T46170	disease resistance
12	54	23.6	3890	2 C89921	hypothetical prote
13	53.5	23.4	350	2 T15178	hypothetical prote
14	53.5	23.4	389	2 D83139	hypothetical prote
15	53	23.1	62	2 D82594	hypothetical prote
16	53	23.1	421	2 B84129	ammonium transport
17	53	23.1	627	2 S14683	Ig mu chain precur
18	52	22.7	116	2 S55542	Ig heavy chain V r
19	52	22.7	146	4 S33905	Ig heavy chain pre
20	52	22.7	214	2 PC4202	monoclonal antibod
21	52	22.7	221	2 S49220	Ig gamma-1 chain -
22	52	22.7	231	2 PC4155	Ig gamma-2b chain
23	52	22.7	1124	2 B45557	HIV-1 retropepsin
24	51.5	22.5	262	2 A10594	probable molybdenu
25	51.5	22.5	282	2 H65116	hypothetical prote
26	51.5	22.5	342	2 D84180	hypothetical prote
27	51.5	22.5	594	2 B90586	transport protein
28	51.5	22.5	1466	2 A85989	hypothetical prote
29	51.5	22.5	1266	2 F91143	hypothetical prote

30	51	22.3	328	2 H75251	glycerol-3-phospha
31	51	22.3	339	2 E95962	probable taurine u
32	51	22.3	372	2 S76427	hypothetical prote
33	51	22.3	393	2 G69178	conserved hypotet
34	51	22.3	423	2 S76384	hypothetical prote
35	51	22.3	662	2 B89875	Quinol oxidase pol
36	51	22.3	865	2 E84718	hypothetical prote
37	50.5	22.1	70	2 AD1647	hypothetical prote
38	50.5	22.1	107	2 A27646	Ig heavy chain V r
39	50.5	22.1	119	2 PH0099	Ig heavy chain V r
40	50.5	22.1	120	2 E49590	Ig heavy chain V r
41	50.5	22.1	173	2 S20003	Ig heavy chain pre
42	50.5	22.1	301	2 C95360	5-dehydro-4-deoxyg
43	50.5	22.1	500	2 C70175	L-lactate permease
44	50.5	22.1	549	2 C83677	L-lactate permease
45	50	21.8	166	2 PL0012	Ig heavy chain pre

ALIGNMENTS

RESULT 1

VCFVAS
env polyprotein - avian spleen necrosis virus
N:Alternate names: coat polyprotein
N:Contains: coat protein gp22; coat protein gp73
C:Species: avian spleen necrosis virus
C>Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C:Accession: A38212
R:Kewairamani, V.N.; Panganiban, A.T.; Emerman, M.
J. Virol. 66, 3026-3031, 1992
A:Title: Spleen necrosis virus, an avian immunosuppressive retrovirus, shares a receptor.
A:Reference number: A38212; MUID:92219390; PMID:1313915
A:Accession: A38212
A:Molecule type: DNA
A:Residues: 1-567 <K&W>
A:Cross-references: UNIPROT:P31796; GB:M87666
C:Genetics:
A:Gene: env
C:Superfamily: type C retrovirus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F:1-401/Domain: extracellular #status predicted <EXT>
F:1-397/Product: coat protein gp73 #status predicted <CP1>
F:363-379/Region: hydrophobic #status predicted
F:394-397/Region: cleavage processing #status predicted
F:398-567/Product: coat protein gp22 #status predicted <CP2>
F:402-418/Domain: transmembrane #status predicted <TM1>
F:419-567/Domain: intracellular #status predicted <INT>
F:245,274,306,328,335/Binding site: carbohydrate (Asn) #status predicted

Query Match 79.9%; Score 183; DB 1; Length 567;
Best Local Similarity 100.0%; Pred. No. 5.8e-17; Mismatches 0; Indels 0; Gaps 0;
Matches 35; Conservative 0;

Qy 1 MDCLTNLSAEGKVDQASKILILLVAVMGFGTTAE 35
|||||
Db 1 MDCLTNLSAEGKVDQASKILILLVAVMGFGTTAE 35

RESULT 2

VCVDFAR
env polyprotein - avian reticuloendotheliosis virus
N:Alternate names: coat polyprotein
N:Contains: coat protein gp22; coat protein gp73
C:Species: avian reticuloendotheliosis virus
C>Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 09-Jul-2004
C:Accession: A03999
R:Wilhelmsen, K.C.; Eggleston, K.; Temin, H.M.
J. Virol. 52, 172-182, 1984
A:Title: Nucleotide acid sequences of the oncogene v-rel in reticuloendotheliosis virus
A:Reference number: A93003; MUID:85009850; PMID:6090694
A:Accession: A03999
A:Molecule type: DNA

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AE0189
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-331 <KUR>
A;Cross-references: UNIPROT:Q82FX0; GB:AL590842; PIDN:CAC90376.1; PID:g15979596; GSPDB:G
C;Genetics:
A;Gene: YPO1553
C;Superfamily: 1-arabinose transport system permease arah

Query Match 24.0%; Score 55; DB 2; Length 331;
Best Local Similarity 40.0%; Pred. No. 10;
Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 17 ASKILLVAVWVGFTTAEV 36
| : :: | | : |||| :
Db 87 ANAVVVSLSFYVYGGTGTGAI 106
| : :: | | : |||| :
RESULT 8
S37483
IG gamma-2a chain - mouse
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C;Accession: S37483
R;Ducancel, F.F.D.
submitted to the EMBL Data Library, February 1993
A;Reference number: S37483
A;Accession: S37483
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-469 <DUC>
A;Cross-references: EMBL:X70423; NID:g406252; PIDN:CAA49868.1; PID:g406253
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: immunoglobulin
F;276-345/Domain: immunoglobulin homology <IMM>

Query Match 24.0%; Score 55; DB 2; Length 469;
Best Local Similarity 45.5%; Pred. No. 14;
Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 24 LVAVWVGFTTAEVSTARAQPA 45
| : :: | | : |||| :
Db 125 LLDYWGQGITLTVSSAKTTAPS 146
| : :: | | : |||| :
RESULT 9
G83692
hypothetical protein BH0343 [imported] - *Bacillus halodurans* (strain C-125)
C;Species: *Bacillus halodurans*
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: G83692
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: G83692
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-214 <STO>
A;Cross-references: UNIPROT:Q9KFX7; GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BA040
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH0343

Query Match 23.6%; Score 54; DB 2; Length 214;
Best Local Similarity 34.3%; Pred. No. 8;
Matches 12; Conservative 6; Mismatches 7; Indels 10; Gaps 1;

QY 14 VDOASKILL-----VAVWVGFTTAEVST 38
| : :: | | : |||| :
Db 36 IRKASLLFLLFTGITALAALAWLYNKTGBIST 70
| : :: | | : |||| :
RESULT 10
AI2012
hypothetical protein alr1655 [imported] - *Nostoc* sp. (strain PCC 7120)
C;Species: *Nostoc* sp. PCC 7120
A;Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AI2012
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AI2012
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-510 <KUR>
A;Cross-references: UNIPROT:Q8YWF8; GB:BA000019; PIDN:BA078021.1; PID:g17135475; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr1655

Query Match 23.6%; Score 54; DB 2; Length 510;
Best Local Similarity 35.7%; Pred. No. 21;
Matches 15; Conservative 10; Mismatches 13; Indels 4; Gaps 1;

QY 2 DCLTNLSAEKGVDOASKILLVAVWVGFTTAEVSTARAQ 43
| : :: | | : |||| :
Db 234 EAITDVSQARVKQAQSLDELLA---GTRSEVITAQAR 271
| : :: | | : |||| :
RESULT 11
T46170
disease resistance-like protein - *Arabidopsis thaliana*
N;Alternate names: protein T4D2.170
C;Species: *Arabidopsis thaliana* (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46170
R;Nykatuna, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.;
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z23025
A;Accession: T46170
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-904 <NYA>
A;Cross-references: UNIPROT:Q9SCN7; EMBL:AL132958
A;Experimental source: cultivar Columbia; BAC clone T4D2
C;Genetics:
A;Map position: 3
A;Introns: 72/3; 92/3; 142/1; 167/1; 186/1
A;Note: T4D2.170

Query Match 23.6%; Score 54; DB 2; Length 904;
Best Local Similarity 47.4%; Pred. No. 37;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 15 DQASKILLVAVWVGFTT 33
| : :: | | : |||| :
Db 846 DDESGLLDIVLWWSLGT 864
| : :: | | : |||| :
RESULT 12
C89921
hypothetical protein ehbB [imported] - *Staphylococcus aureus* (strain N315)
C;Species: *Staphylococcus aureus*
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: C89921
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 12:13:31 ; Search time 93 Seconds
(without alignments)
247.780 Million cell updates/sec

Title: US-10-089-278-6_COPY_1_45
Perfect score: 229
Sequence: 1 MDCLTLNRSABGKVDQASKI.....AMWGFGTTAEVSTARAQA 45

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	183	79.9	567	1 ENV_AVISN	P31796 avian splee
2	170	74.2	586	2 Q91G02	Q91G02 fowlpox vir
3	161	70.3	582	1 ENV_AVIRE	P03399 avian retic
4	65.5	28.6	2020	2 Q8ST27	Q8ST27 dictyosteli
5	65.5	28.6	2481	2 Q9FR53	Q9FR53 arabidopsis
6	65.5	28.6	2513	2 Q9LPM4	Q9LPM4 arabidopsis
7	63	27.5	323	2 Q65FX1	Q65FX1 bacillus li
8	61.5	26.9	473	2 Q91Z05	Q91Z05 mus musculu
9	59.5	26.0	651	2 Q8KNF9	Q8KNF9 micromonosp
10	58.5	25.5	1031	2 Q8ZM07	Q8ZM07 salmonella
11	58	25.3	176	2 Q8Y1D9	Q8Y1D9 ralestonia s
12	58	25.3	527	2 Q8GY76	Q8GY76 arabidopsis
13	58	25.3	538	2 Q64683	Q64683 arabidopsis
14	57	24.9	276	2 Q8HEH5	Q8HEH5 citharus li
15	57	24.9	364	2 Q7MHZ1	Q7MHZ1 vibrio vuln
16	57	24.9	364	2 Q8DBX6	Q8DBX6 vibrio vuln
17	57	24.9	459	2 Q7NNF6	Q7NNF6 gloeobacter
18	57	24.9	470	2 Q82282	Q82282 chlamydophi
19	56.5	24.7	247	2 Q93KB4	Q93KB4 erwina chr
20	56.5	24.7	272	2 Q75CF2	Q75CF2 ashbya goss
21	55.5	24.2	364	2 Q8PBC5	Q8PBC5 xanthomonas
22	55	24.0	171	2 Q9PV14	Q9PV14 soybean mos
23	55	24.0	193	2 Q9PVJ7	Q9PVJ7 soybean mos
24	55	24.0	194	2 Q9PVI9	Q9PVI9 soybean mos
25	55	24.0	331	2 Q66C45	Q66C45 yersinia ps
26	55	24.0	331	2 Q82FX0	Q82FX0 yersinia ps
27	54	23.6	142	2 Q88GV3	Q88GV3 pseudomonas
28	54	23.6	214	2 Q9KFX7	Q9KFX7 bacillus ha
29	54	23.6	282	1 MODD_PASMU	Q9CLU4 pasteurella
30	54	23.6	381	2 Q94RJ0	Q94RJ0 chinaera mo
31	54	23.6	460	2 Q7QJ04	Q7QJ04 anopheles g

32	54	23.6	495	2	Q7UFJ4	Q7UFJ4 rhodopirell
33	54	23.6	510	2	Q8YWF8	Q8YWF8 anabaena sp
34	54	23.6	870	2	Q6L1Z4	Q6L1Z4 picorhizus
35	54	23.6	904	2	Q9SCN7	Q9SCN7 arabidopsis
36	54	23.6	1240	2	Q706G7	Q706G7 anopheles g
37	54	23.6	1263	2	Q6ATH2	Q6ATH2 oryza sativ
38	54	23.6	3890	2	Q99U53	Q99U53 staphylococ
39	54	23.6	3890	2	Q7ASM1	Q7ASM1 staphylococ
40	54	23.6	9904	2	Q8NMQ6	Q8NMQ6 staphylococ
41	53.5	23.4	169	2	Q8R9L0	Q8R9L0 paracoccus
42	53.5	23.4	350	2	Q81G65	Q81G65 caenorhabdi
43	53.5	23.4	389	2	Q9HWY6	Q9HWY6 pseudomonas
44	53.5	23.4	441	2	Q9K4A3	Q9K4A3 streptomyce
45	53.5	23.4	801	2	Q7XEM7	Q7XEM7 oryza sativ

ALIGNMENTS

RESULT 1

ID	ENV_AVISN	STANDARD;	PRT;	567 AA.
AC	P31796;			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Env polyprotein (Coat polyprotein) [Contains: Coat protein GP73; Coat protein GP22].			
DE	Name=ENV;			
OS	Avian spleen necrosis virus.			
OC	Viruses; Retrovirdae; Retroviridae; Avian type C retroviruses.			
OX	NCBI_TaxID=11899;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Clone PB101;			
RX	MEDLINE=92219390; PubMed=1313915;			
RX	Kewalramani V.N., Panganiban A.T., Emerman M.;			
RT	"Spleen necrosis virus, an avian immunosuppressive retrovirus, shares a receptor with the type D simian retroviruses.";			
RL	J. Virol. 66:3026-3031(1992).			
CC	!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.			

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	EMBL; M37666; ; NOT ANNOTATED_CDS.
DR	PIR; A38212; VCFVAS.
DR	HSSP; P03385; LMOP.
DR	InterPro; IPR008965; Cellul_bind.
DR	InterPro; IPR002050; Env_polyprotein.
DR	Pfam; PF00429; Tlv_coat; 1.
FT	Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT	CHAIN 1 397
FT	CHAIN 398 567
FT	TRANSMEM 402 418
FT	CARBOHYD 245 245
FT	CARBOHYD 280 274
FT	CARBOHYD 280 280
FT	CARBOHYD 306 306
FT	CARBOHYD 319 319
FT	CARBOHYD 328 328
FT	CARBOHYD 335 335
FT	CARBOHYD 491 491
SQ	SEQUENCE 567 AA; 61596 MW; 204EA57C32159175 CRC64;

Query Match 79.9%; Score 183; DB 1; Length 567;
Best Local Similarity 100.0%; Pred. No. 4.7e-17;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MDCLTNLSAEGKVDQASKILLVAVWVGFTTAE 35
DB 1 MDCLTNLSAEGKVDQASKILLVAVWVGFTTAE 35

RESULT 2
ID Q9IGU2 PRELIMINARY; PRT; 586 AA.
AC Q9IGU2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Envelope glycoprotein.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RA Singh P., Kim T.-J., Tripathy D.N.;
RT "Re-emerging fowlpox: evaluation of isolates from vaccinated flocks.";
RL Avian Pathol. 29:449-455(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Singh P., Schitzlein W.M., Tripathy D.N.;
RT "Reticuloendotheliosis Virus Sequences Within the Genomes of Field
RT Strains of Fowlpox Virus Display Variability.";
RL J. Virol. 0:0-0(2003).
RN [3]
RP SEQUENCE FROM N.A.
RA Schmitzlein W.M., Singh P., Srinivasan V., Tripathy D.N.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF246698; AAF81698.2; -.
DR HSP; P03385; IMOP.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002050; Env_polyprotein.
DR Pfam; PF00429; TLV_coat; 1.
KW Envelope protein.
SQ SEQUENCE 586 AA; 63723 MW; DBB221FA9FC562C3 CRC64;

Query Match 74.2%; Score 170; DB 2; Length 586;
Best Local Similarity 91.4%; Pred. No. 3.4e-15;
Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MDCLTNLSAEGKVDQASKILLVAVWVGFTTAE 35
DB 1 MDCLTNLSAEGKVDQASKILLVAVWVGFTTAE 35

RESULT 3
ENV_AVIRE
ID ENV_AVIRE STANDARD; PRT; 582 AA.
AC P03399;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Env polyprotein (Coat polyprotein) [Contains: Coat protein GP73; Coat
DE protein GP22].
GN Name=ENV;
OS Avian reticuloendotheliosis virus.
OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.
OX NCBI_TaxID=11636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A;
RX MEDLINE=8509850; PubMed=6090694;
RA Wilhelmson K.C., Eggleston K., Temin H.M.;
RT "Nucleic acid sequences of the oncogene v-rel in reticuloendotheliosis
RT virus strain T and its cellular homolog, the proto-oncogene c-rel.";
RL J. Virol. 52:172-182(1984).

```

```

CC -|- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -|- MISCELLANEOUS: Strain A is a helper virus of the strain T.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X01455; CAA25686.1; -.
DR PIR; A03999; VCVDAAR.
DR HSP; P03385; IMOP.
DR InterPro; IPR002050; Env_polyprotein.
DR Pfam; PF00429; TLV_coat; 1.
KW Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT CHAIN 1 391
FT CHAIN 392 582
FT CARBOHYD 241 241
FT CARBOHYD 301 301
FT CARBOHYD 314 314
FT CARBOHYD 485 485
FT CARBOHYD 485 485
SQ SEQUENCE 582 AA; 64138 MW; CD2560ADF0C26D32 CRC64;

Query Match 70.3%; Score 161; DB 1; Length 582;
Best Local Similarity 85.7%; Pred. No. 6.4e-14;
Matches 30; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDCLTNLSAEGKVDQASKILLVAVWVGFTTAE 35
DB 1 MDCLTNLSAEGKVDQASKILLVAVWVGFTTAE 35

RESULT 4
Q8ST27 PRELIMINARY; PRT; 2020 AA.
AC Q8ST27;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to Dictyostelium discoideum (slime mold). CIGB protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44889;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tungal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC114265; AAL86966.2; -.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR008615; FNIP.
DR InterPro; IPR001611; LRR.
DR Pfam; PF05725; FNIP; 6.
DR Pfam; PF00560; LRR_1; 3.
DR SMART; SM00248; ANK; 1.
DR SEQUENCE 2020 AA; 232614 MW; 01E221CE63114DD3 CRC64;

Query Match 28.6%; Score 65.5; DB 2; Length 2020;
Best Local Similarity 46.7%; Pred. No. 8.6;
Matches 14; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

QY 2 DCILTNLSAEGKVDQASKILLVAVWVGFTTAE 30

```

Db	686	DCISNSNSEKKTKDQVSKYILLIKSGFG	715
RESULT 5			
Q9FR53		PRELIMINARY;	PRT; 2481 AA.
AC	Q9FR53		
DT	01-MAR-2001	(TrEMBLrel. 16, Created)	
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	PTOR.		
GN	Name=TOR;		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta;		
OC	Spermatophyta; Magnoliophyta; Eudicotyledons; rosids;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=21980689; PubMed=11983923; DOI=10.1073/pnas.092141899;		
RA	Menand B., Desnos T., Nussaume L., Berger F., Bouchez D., Meyer C.,		
RA	Robaglia C.;		
RT	"Expression and disruption of the Arabidopsis TOR (target of rapamycin) gene.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:6422-6427(2002).		
DR	EMBL; AF178967; AAG43423.1; -.		
DR	HSSP; P42345; 1FAP.		
DR	GO; GO:0016773; P:phosphotransferase activity, alcohol group . . . ; IEA.		
DR	InterPro; IPR008938; ARM.		
DR	InterPro; IPR003151; FAT.		
DR	InterPro; IPR009076; FRAP_FKBP12_bind.		
DR	InterPro; IPR000357; HEAT.		
DR	InterPro; IPR011009; Kinase like.		
DR	InterPro; IPR004043; PI3_Pi4_kinase.		
DR	InterPro; IPR008941; TPR-like.		
DR	Pfam; PF02259; FAT; 1.		
DR	Pfam; PF02260; FATC; 1.		
DR	Pfam; PF02985; HEAT; 9.		
DR	Pfam; PF00454; PI3_Pi4_kinase; 1.		
DR	SMART; SM00146; PI3K; 1.		
DR	PROSITE; PS00915; PI3_4_KINASE_1; 1.		
DR	PROSITE; PS00916; PI3_4_KINASE_2; 1.		
DR	PROSITE; PS0290; PI3_4_KINASE_3; 1.		
SQ	SEQUENCE 2481 AA; 279187 MW; DA663EA9A9366F93 CRC64;		
Query Match	28.6%; Score 65.5; DB 2; Length 2481;		
Best Local Similarity	43.2%; Pred. No. 11;		
Matches	16; Conservative 3; Mismatches 15; Indels 3; Gaps 1;		
Qy	3	CLTNLSAEGKVDQASKILLLVAVWVGFGTTAEVSTA 39	
Db	1794	CAAN---AKGVDSLQDILRLTLFWNHGATADVQTA 1827	
RESULT 6			
Q9LPM4		PRELIMINARY;	PRT; 2513 AA.
AC	Q9LPM4		
DT	01-OCT-2000	(TrEMBLrel. 15, Created)	
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	F2110.9 protein.		
GN	Name=f2110.9;		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta;		
OC	Spermatophyta; Magnoliophyta; Eudicotyledons; rosids;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Sakano H., Liu S.X., Yu G., Lee J., Lenz C., Pham P., Toriumi M.,		

Db	686	DCISNSNSEKKTKDQVSKYILLIKSGFG	715
RESULT 5			
Q9FR53		PRELIMINARY;	PRT; 2481 AA.
AC	Q9FR53		
DT	01-MAR-2001	(TrEMBLrel. 16, Created)	
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	PTOR.		
GN	Name=TOR;		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta;		
OC	Spermatophyta; Magnoliophyta; Eudicotyledons; rosids;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21980689; PubMed=11983923; DOI=10.1073/pnas.092141899;		
RA	Menand B., Desnos T., Nussaume L., Berger F., Bouchez D., Meyer C.,		
RA	Robaglia C.;		
RT	"Expression and disruption of the Arabidopsis TOR (target of rapamycin) gene.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:6422-6427(2002).		
DR	EMBL; AF178967; AAG43423.1; -.		
DR	HSSP; P42345; 1FAP.		
DR	GO; GO:0016773; P:phosphotransferase activity, alcohol group . . . ; IEA.		
DR	InterPro; IPR008938; ARM.		
DR	InterPro; IPR003151; FATC.		
DR	InterPro; IPR009076; FRAP_FKBP12_bind.		
DR	InterPro; IPR000357; HEAT.		
DR	InterPro; IPR011009; Kinase like.		
DR	InterPro; IPR004043; PI3_Pi4_kinase.		
DR	InterPro; IPR008941; TPR-like.		
DR	Pfam; PF02259; FAT; 1.		
DR	Pfam; PF02260; FATC; 1.		
DR	Pfam; PF02985; HEAT; 9.		
DR	Pfam; PF00454; PI3_Pi4_kinase; 1.		
DR	SMART; SM00146; PI3K; 1.		
DR	PROSITE; PS00915; PI3_4_KINASE_1; 1.		
DR	PROSITE; PS00916; PI3_4_KINASE_2; 1.		
DR	PROSITE; PS0290; PI3_4_KINASE_3; 1.		
SQ	SEQUENCE 2481 AA; 279187 MW; DA663EA9A9366F93 CRC64;		
Query Match	28.6%; Score 65.5; DB 2; Length 2481;		
Best Local Similarity	43.2%; Pred. No. 11;		
Matches	16; Conservative 3; Mismatches 15; Indels 3; Gaps 1;		
Qy	3	CLTNLSAEGKVDQASKILLLVAVWVGFGTTAEVSTA 39	
Db	1794	CAAN---AKGVDDSLQDILRLTLFNFHGATADVQTA 1827	
RESULT 6			
Q9LPM4		PRELIMINARY;	PRT; 2513 AA.
AC	Q9LPM4		
DT	01-OCT-2000	(TrEMBLrel. 15, Created)	
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	F2110.9 protein.		
GN	Name=f2110.9;		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta;		
OC	Spermatophyta; Magnoliophyta; Eudicotyledons; rosids;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Sakano H., Liu S.X., Yu G., Lee J., Lenz C., Pham P., Toriumi M.,		

Query Match 25.5%; Score 58.5; DB 2; Length 1031;

[illegible]

CC -!- SUBUNIT: The main subunits of complex b-c1 are:

CC -!- SUBUNIT: The main subunits of complex b-c1 are:

Search completed: August 25, 2005, 15:29:38
Job time : 100 secs

QY 4 LTNLRSAGCKVDQASKILI-----LLVAVWVGFT-----TAEVS 37
:||||: |: |: | ||| |:
pb 1 MTNLKTHPLKCVANDAVIDLPAPSNLUSAWNFGSLGLCMLMAOLLTGLELAMHYTADIA 60
:||||: |: |: | ||| |:

Query Match	24.9%;	Score 57;	DB 2;	Length 364;
Best Local Similarity	46.2%;	Pred. No. 23;		
Matches 12:	Conservative	7;	Mismatches	5;
			Indels	2;
			Gaps	1;

Qy 20 ILILLVAWWGFGTTA--EVSTARAAQ 43
|::| | | | |::| | | | |
pb 21 IILGLAWLGFSSKADETSTSKAAQ 46

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☒ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☒ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (uspto)